Fall 11-20-2018

lcp1 mutant zebrafish: A look at neutrophils, cancer, and gene compensation

Taylor Mitchell
DePaul University, taylorannmitchell93@gmail.com

lcp1 mutant zebrafish: A look at neutrophils, cancer, and gene compensation

A Thesis

Presented in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

December, 2018

BY

Taylor A. Mitchell

Department of Biological Sciences

College of Science and Health

DePaul University

Chicago, Illinois

TABLE OF CONTENTS

Taylo
Delet

Acknowledgments

To begin, I would like to deeply thank my advisor Dr. Elizabeth LeClair, who I first met when I was still an undergraduate. She was a mentor to me back then, and being able to have her as an official mentor in the capacity of a research advisor has been phenomenal. Her passion for any field of biology is extremely contagious and helped my interest to increase even more. Her guidance throughout this graduate program has been helpful to not only my growth as a researcher and student, but as a life-long learner as well.

I would also like to show gratitude to my thesis committee members Drs. Windsor Aguirre and Talitha Rajah. Not only have they both guided me in my thesis project they were also instructors of courses I took that helped advanced my knowledge in science, and furthered my understanding of my thesis work. They both have had wonderful input, which has made me a better scientist.

My appreciation is also extended to the biological sciences faculty who I have interacted with as instructors I took classes from, professors I was a teaching assistant for, and the biology staff for encouraging me during these past two years.

Finally, I would like to thank all of the other members of the LeClair lab for their support in planning experiments, bouncing ideas back and forth, and helping me execute my experiments. I would also like to thank my thesis cohort for the constant reminders that we were all on this journey together, and my family for being a constant support system in my educational career, and encouraging me every step of the way.

Abstract

Lymphocyte cytosolic protein 1 (*lcp1* or L-plastin*)* is a small actin-bundling protein that is typically only expressed in motile leukocytes, such as neutrophils and macrophages. However, it is also overexpressed in cancer cells, which may be related to tumor metastasis. Using CRISPR/Cas9 gene editing, our lab has created zebrafish that are genetic knockouts for *lcp1* in order to better understand the relationship between L-plastin and cell motility.

Previous studies on L-plastin knockout mice have shown that the mutants have a decreased immune response, and therefore I predicted that our zebrafish mutants might have impaired development or distribution of immune cells. Other experiments have demonstrated that decreasing *lcp1* in tumor cells in mice decreases the growth, invasiveness, and metastasis of cancer cells both *in vitro* and *in vivo*. Inducing tumors in the zebrafish mutant might show decreased tumor penetrance, which would further confirm the relationship between *lcp1* expression and cancer. Finally, I investigated the entire transcriptome of the *lcp1* mutant zebrafish embryos and compared that to the wildtype transcriptome using RNA-seq. This provides information about any genes that may be differentially expressed in the mutant in order to compensate for the lack of *lcp1* mRNA.

Overall, this study aims to further characterize *lcp1* mutant zebrafish in terms of their distribution of immune cells, susceptibility to tumors, as well as differences in their transcriptome in comparison to wildtype zebrafish. This data can expand on the current knowledge of the role of *lcp1* in hematopoiesis and cancer, as well as adding new understanding to the dynamic between *lcp1* and other genes in the transcriptome in relation to transcriptional adaptation.

Introduction

Eukaryotic cell mechanics

There are two types of cells: prokaryotic and eukaryotic. Eukaryotic cells are large and have a cytoskeleton. The cytoskeleton is essential for the organization of organelles within the cell, allows for cell movement, and connects the cell to the extracellular matrix. Defects of the cytoskeleton could lead to adverse effects such as abnormal movement or morphology. For example, mutations in α-actin have been linked to congenital myopathy in infants, which can lead to difficulty breathing or feeding (Clarkson et al. 2004).

Components of the cytoskeleton

The cytoskeleton has three major components: microtubules, intermediate filaments, and actin (Fletcher and Mullins 2010). Each component has a specific architecture and function. Microtubules are long and hollow with the largest diameter, (24-25 nm), of the cytoskeletal molecules. They are important for cell division as they form the mitotic spindle to organize chromosomes. Intermediate filaments are slightly smaller in diameter than microtubules (8-11 nm). These surround the nucleus and extend to the plasma membrane. Actin filaments are the thinnest cytoskeletal elements (5-9 nm)**,** and are best known for their role in muscle contraction, in conjunction with myosin (Schliwa 1986). However, many proteins can bind actin for various functions within eukaryotic cells (**Fig. 1).**

In 1942, Bruno Ferenc Straub first recognized actin as a separate protein from myosin by exposing ground muscle tissue to high salt concentrations (Szent-Györgyi 2004). A short exposure to salt allowed the extraction of a low viscosity protein, then called ‘myosin A’, and an overnight exposure led to a high viscosity protein, then called ‘myosin B’. Myosin B was renamed ‘actin’ by Straub’s advisor in that same year. Straub also was the first to observe that

actin was present in two forms: globular actin (G-actin) and filamentous actin (F-actin). However the amino acid sequence of G-actin was not confirmed until 1973, using rabbit muscle (Elzinga et al. 1973).

G-actin monomers can polymerize spontaneously into F-actin filaments, although this process is slow due to the activation of the monomers needed to make the reaction more kinetically favorable (Cooper et al. 1983). In living cells, actin polymerization is highly regulated, and many studies have shown the important roles of actin-polymerizing and bundling proteins (Paul and Pollard 2009). These helper proteins change the structure of the cytoskeleton and vary by the cell type, depending on the need of the cell to expand and contract (Otto 1994). The modification of the actin cytoskeleton is especially important in motile cells, such as osteoclasts (Ma et al. 2010) and white blood cells, also known as leukocytes (Boxer et al. 1976). **Figure 2** shows several lineages of leukocytes, including macrophages, neutrophils, T-cells, B-cells and mast cells. When crawling, these cells use actin-rich microspikes and/or lamellipodia that project from the plasma membrane (Bray 1992) **(Fig. 3B).**

Actin-bundling proteins: the plastin family

As stated previously, crawling cells contain actin that, when assembled and dissembled in the proper direction and with the correct rate, allows the cell to move in their characteristic motion. However, the actin must be properly bundled to carry out its function. There are many proteins that can bundle actin (Amos and Amos 1991), but one key group is the plastin protein family. Plastins are present in all eukaryotes, and are highly conserved in sequence and function (Arpin et al. 1994). In vertebrates, there are 3 family members: T-plastin, which is expressed in the basement membrane (Dor-On et al. 2017); I-plastin, which is expressed in the intestines and

kidneys (Lin et al. 1994), and L-plastin, restricted to leukocytes (Lin et al. 1988). Although expressed in different tissues, the plastin proteins share similar actin-binding domains, and share regulatory elements. It is known that the ability to bind to actin is dependent upon Ca^{2+} (Amos and Amos 1991). In humans, it was found that L-plastin's first actin-binding domain will bind to actin regardless of Ca^{2+} concentration, whereas the second actin-binding domain has a decreased affinity for actin in higher Ca^{2+} concentrations (Schwebach et al. 2017). This confirms that Ca^{2+} is a regulatory element for the plastin proteins. The other known regulatory element of the plastins is phosphorylation—specifically, phosphorylation of one or more N-terminal serines in the amino acid chain. This has been investigated by Janji et al. 2006, who showed that *lcp1* that had a Ser5Ala substitution and could not be phosphorylated reduced the protein's localization to membrane extensions (Janji et al. 2006).

L-Plastin, an immune-specific actin bundler

Of the three plastin family members, two have rather broad expression (I-plastin and T-plastin). However the third is far more specific. Lymphocyte cytosolic protein 1 (LCP1 or L-plastin) was first described in 1985 when it was detected in transformed human fibroblast cell cultures (Goldstein et al. 1985). LCP1 was further characterized as a leukocyte-specific expressed gene by cDNA analysis leukocytes (Lin et al. 1988). Subsequently, LCP1 was confirmed as one of the highest expressed genes in leukocytes, regardless of myeloid or lymphoid branches (Hashimoto et al. 2003).

LCP1 is a small protein (624 amino acids) that consists of two Ca^{2+} binding domains, two actin-binding domains, and a phosphorylation site at the N terminus (Shinomiya 2012) **(Fig. 3A).** Each molecule of L-plastin can bind two adjacent actin filaments. When one actin filament attaches to L-plastin, it induces a conformational change to activate more actin binding,

as previously mentioned (Galkin et al. 2008). The ability of L-plastin to bind to actin is regulated by intracellular calcium, as previously described by Schwebach et al, 2017, and serine phosphorylation (Janji et al. 2006). This was described *in vivo* by using wild type and LPC-1 mutant cell lines of mice; lines that had phosphorylated LCP1 showed increased actin binding (Morley 2013).

L-plastin expression and function in immune cells

LCP1 has a limited expression pattern, and is typically only expressed in cells of the leukocyte lineage, such as neutrophils **(Fig. 2**) (Lin, Chen, et al. 1993; Lin, Park, et al. 1993, 1993). Because actin is one of the most abundant proteins of the cytoskeleton, most actin-binding proteins are present in almost all eukaryotic cells as "housekeeping genes"; however, L-plastin is found exclusively in immune cells such as neutrophils and macrophages and is often used as a marker for these cell types (Mathias et al. 2009). This suggests that L-plastin has an important role in immune cell function.

An important milestone in L-plastin research has been construction of an L-plastin knockout mouse (LPL -/-) (Chen et al. 2003). Although these mice are viable and fertile, they are susceptible to certain infections. For example, homozygous mutant mice and wild-type littermates have similar numbers of polymorphonuclear neutrophils (PMN), but mutant mice were unable to kill off incoming *Staphylococcus aureus* (Chen et al. 2003). *In vitro,* cell lines derived from these mice showed decreased activation of SYK, which is a tyrosine kinase required for the 'respiratory burst' needed to kill an infection such as *S. aureus,* in Western blots. Thus, it is possible that LCP1 deficiency makes neutrophils less able to combat infections.

L-plastin deficiency has also been shown to affect other types of immune cells. The T-cells in these mice are poorly activated as shown by proliferation assays, and antigen-specific

responses, as shown by inflammatory response expression (Wang et al. 2010). Furthermore, these LCP1 deficient mice have defects in germinal center (GC) development, which is needed for the long-term production of B-cells (Todd et al. 2013). Finally, the macrophages of LCP1 deficient mice have also been studied. When subjected to an inhaled bacterial challenge from *Streptococcus pneumoniae,* could not clear the infection from their lung alveoli and died at alarming rates (~80% mortality, compared to 8% in wild types). Using a flow cytometer for a marker specific to macrophages, it was determined that LCP1 deficient mice had fewer macrophages in the alveoli. However, by using fluorescein- labeled bacteria, it was confirmed that the few macrophages that were present could ingest the bacteria. It was concluded that the LPL -/- mutant macrophages are phagocytosis-competent but migration-deficient, and cannot reach the site of infection in sufficient numbers (Todd et al. 2016).

Furthermore, LCP1 specifically promotes podosome longevity in relation to motility. A podosome, or "foot organ", is an adhesive cellular structure used by macrophages during migration. Macrophages from wild type (WT) mice and LCP1-deficient mice were isolated and examined using immunohistochemistry and time-lapse microscopy *in vitro* (Zhou et al. 2016). In WT macrophages F-actin and LCP1 were co-localized in the podosomes. LCP1-deficient macrophages still had podosomes, but the stability of the structure was decreased as the podosomes persisted for a much shorter amount of time. Overall, it is clear that LCP1 has an important role in mouse immune cells studied *in vivo* as well as *in vitro*. However, there is still much to learn about the function of LCP1 in an organism's immune response.

L-plastin expression and function in cancer cells

As previously stated, L-plastin is typically highly restricted to normal immune cells; however, numerous studies have shown increased expression of L-plastin in non-immune derived cancer cell lines (breast, colon, kidney, etc.) and primary tumors **(Fig. 4).** For example, Park et al. assayed twelve cancer cell lines, eight diploid human cell lines, and four transformed human fibroblast cell lines using Northern blotting and reverse-transcription polymerase chain reaction (Park et al. 1994). Of the 48 neoplastic cell line samples tested, 92% showed some level of LCP1 expression, with some of the highest levels detected in hematopoietic tumors, cervical carcinoma and colon carcinoma. These observations indicate that L-plastin is abnormally activated in cancer cells, by some unknown mechanism. It is therefore likely that the cancer cells are utilizing L-plastin to bundle actin for movement, and that manipulating L-plastin can modify cancer cell behavior (Shinomiya 2012).

Multiple studies have since confirmed that LCP1 expression can be used as a marker for primary tumors. Using cDNA samples from primary and metastatic colorectal tumors from the same patient, it was found that LCP1 was expressed at roughly twice as much in the metastatic tumor rather than the primary tumor (Otsuka et al. 2001). Immunohistochemistry for LCP1 was then performed on 58 colon tumors varying from Stage II to Stage IV to better understand the correlation between LCP1 expression and tumor progression. LCP1 was moderately to strongly stained in 78.6% of Stage III and IV tumors, but only 28.5% of Stage II tumors. This suggests that the expression level of LCP1 is correlated with tumor stage and could be a biomarker for metastasis. A similar experiment was conducted on urinary bladder tumor samples, which

showed that LCP1 had significant expression associated with tumor grade (p=0.035) and growth pattern (p=0.047).

LCP1 can also be used as a biomarker in ovarian tumors, and has an interesting expression pattern. Using matrix-associated laser desorption/ionization-mass spectrometry (MALDI-IMS) technology, and confirmed by Western Blot, immunofluorescence, and immunohistochemistry, LCP1 was found to be highly upregulated in the interface zone (IZ), or the region between front of the tumor invasion and the normal tissue area (Kang et al. 2010). This is an important observation, further confirming that LCP1 is involved in the invasion process. Finally, LCP1 is highly expressed in oral squamous cell carcinomas (OSCCs). Using qRT-PCR and immunoblot analysis comparing normal human oral keratinocytes to OSCC-derived cell lines, LCP1 was significantly upregulated in the tumor cell derived lines (Koide et al. 2017).

While the presence of LCP1 in multiple tumor cell types has been well described, additional evidence for a causal role for L-plastin in tumor metastasis and invasion comes from *in vitro* and *in vivo* studies that manipulate L-plastin levels. For example, by injecting antisense LCP1 vectors into prostate carcinoma cell lines, Western Blots confirmed that LCP1 expression levels were decreased (Zheng et al. 1999). These cells were then tested for migration and invasion using *in vitro* and *in vivo* assays. *In vitro*, the modified cells were placed in a Matrigel membrane matrix and allowed 24 hours to migrate through the pores. LCP1 knockdown cell lines showed 68%-89% reduction in invasiveness when compared to cancer cell lines with unmodified LCP1. *In vivo*, nude mice were injected with the knock-down tumor lines and 21 days after inoculation, the diaphragm of the mice was analyzed for tumors and scored for invasion based on basement membrane penetration. Mice injected with knock-down cell lines

saw 70-80% reduction in diaphragm muscle invasion. In addition, knocking down L-plastin in human melanoma cell lines using short interfering RNA (siRNA) led to reduced cell migration *in vitro* (Klemke et al. 2007). Bosseler et al. demonstrated that in multiple myeloma cell cultures from patients LCP1 is involved in multiple myeloma drug resistance by analyzing the expression level in LCP1 in drug resistant cell lines and cell lines sensitive to multiple myeloma drug treatments. Drug resistant cell lines showed a higher level of LCP1 expression (Bosseler et al. 2018).

In vivo, a key study showed that mice injected with human prostate cancer cells that ectopically expressed of L-plastin increased the invasion of these tumor cells (Riplinger et al. 2014). Targeting LCP1 or its promoter in human bladder and ovarian cancer tumors was done by using an adenoviral vector with a truncated human LCP1 promoter and cytosine deaminase (CD). CD converts 5-fluorocytosine to 5-fluorouracil, which is toxic to most cells. These vectors were injected into the human cancer cell lines, and then the cell cultures were injected into nude mice. Twenty-one days after injection, mice were analyzed for tumor growth. Compared to the control mice, in which 100% of the mice injected had tumors, none of the mice injected with LCP1 CD vector developed tumors. This suggests that perhaps LCP1 can be used as a target for tumor treatment as the only cells in the body that should expression LCP1 are mature leukocytes and tumor cells (Peng et al. 2001).

Overall, numerous descriptive and experimental studies have shown that L-plastin is abnormally expressed in cancer cells that are not derived from the immune system and that, in some assays, knocking down L-plastin leads to reduced cell motility, which is related to invasiveness and metastasis.

Research Objectives

Zebrafish lcp1 mutants

Our lab has created multiple lines of zebrafish knockouts of L-plastin using CRISPR/Cas9 gene editing (Kell et al. 2018). The *lcp1* gene is located on chromosome 9 in the zebrafish genome and the induced mutations in our lines are within exon 2 of this locus **(Fig. 5 A & B)**. My work focused on the FOXTROT (FX) line of mutants, which have a net 5 base pair (bp) deletion, causing a premature stop codon upstream of the actin-binding domains **(Fig. 5C)**. The lab has described all three mutants lines and confirmed, using Western blots, that none express LCP1 protein either as embryos or as adults **(Fig. 6)** (Kell et al. 2018)**.** Heterozygous and homozygous mutant fish can be distinguished from each other using genotyping PCR, and all genotypes of zebrafish are viable and can reproduce. My work focuses on studying the mutants' immune system and response to tumor-inducing plasmids to further characterize these fish lines.

Zebrafish as a model for the innate immune system

Although the zebrafish immune system has been studied extensively, there is no published research on zebrafish L-plastin mutants and their immune systems. Unlike mice, zebrafish are transparent, which makes the zebrafish an ideal model organism to study immune cell development, migration and behavior (Traver et al. 2003; Trede et al. 2004). Zebrafish neutrophils and macrophages have been well described, and can be easily identified 1-2 days after fertilization.

One of the earliest studies of immunity in this organism used video microscopy to show that, at 25 hours post fertilization (25 hpf), the bloodstream carried motile cells that had phagocytic capabilities (Herbomel et al. 1999). These cells were early macrophages. Several years later, researchers sought to find a screening strategy for macrophages, and they found through *in situ* hybridization that L-plastin mRNA was the best marker for this cell type, to differentiate them from granulocytic cell populations (Crowhurst et al. 2002). More recently, another group used photoactivatable tracers *in vivo* live larvae to observe several lineages of immune cells during zebrafish development. Using an antibody against L-plastin, they determined that primitive macrophages and neutrophils both express this protein in fixed samples (Le Guyader et al. 2008).

These investigations have shown that both neutrophils and macrophages are present in zebrafish at 2dpf, and both cell types express L-plastin. However, there are no studies in zebrafish on the effects of knocking out this immune-specific protein.

Zebrafish as a model for cancer

Zebrafish are used extensively as models for many human diseases including cancer initiation, progression and treatment (Lieschke and Currie 2007; Mione and Trede 2010). Although there have been many different approaches to induce tumors in zebrafish, the two main methods are chemical induction or genetic induction. Chemically, scientists have used dimethylbenzanthracene to induce intestinal cancer, diethylnitrosamine to induce pancreatic carcinoma and N-ethyl-N-nitrosourea to induce testicular cancer (Mizgirev and Revskoy 2010; Mirbahai et al. 2011; Basten et al. 2013). Using chemical models on zebrafish is attractive because the drugs can added to the water directly, many chemicals can be tested at once, and large numbers of zebrafish are fairly easy and cost effective to maintain (Huiting et al. 2015).

However, in order to study specific tumor types, genetically modifying the zebrafish genome is the most convenient option. Although genetic modification involves constructing a plasmid with a human oncogene and a zebrafish promoter, then microinjecting this into the zebrafish at the one-cell stage, these lines can be created to model various different tumor types including leukemia, melanoma, and rhabdomyosarcoma (Yen et al. 2014). With some of these cancer types, a fluorescent label can be used to identify the cancerous cells *in vivo*. For example, the Langenau lab created two separate plasmids to induce rhabdomyosarcoma (RMS), a type of muscle cancer, and simultaneously label the RMS cells *in vivo.* In this system, an oncogenic plasmid, which induces tumors, is co-injected with a reporter plasmid, which makes the tumors light up (Chen and Langenau 2011).

lcp1 zebrafish mutants as a model for transcriptional adaptation

Transcriptional adaptation, sometimes called genetic compensation, is the concept of a biological system overcoming a setback, such as non-functional protein created by CRISPR/Cas9, by increasing the transcription of a different gene (Grether 2005). As described in a review by Darius Balciunas, this process could be one of the explanations for why a mutant does not have a phenotype very different from the wildtype (Balciunas 2018). The first example of zebrafish transcriptional adaptation came from analyzing morphants and mutants for *EGF-like-domain, multiple 7 (egfl7)* and comparing them to wildtype fish. While the morpholino fish had very severe vascular defects, the mutants did not express any visible phenotype. By comparing the transcriptome and proteome of the morpholinos and mutants using RNA profiling and qPCR, and SDS-PAGE and mass spectrometry respectively, it was found that the mutant fish were expressing a set of genes at an upregulated level. Some of these genes were then injected into the morpholinos, which rescued them from their severe phenotype (Rossi et al. 2015).

Due to the lack of L-plastin in our mutants, it is possible that a related gene such as other isoforms of plastin or fimbrins, may have increased expression to compensate as previous studies on other zebrafish knockouts have described (Rossi et al. 2015). It has also shown that LCP1 is expressed in the enveloping layer during gastrulation in wildtype fish, which is likely because the enveloping layer must crawl in a similar manner to leukocyte cells over the deep layer during gastrulation (Kimmel et al. 1995). *Lcp1* may be expressed for motility of these cells. However, our mutant zebrafish develop through gastrulation and are viable without their EVL cells expressing the gene.

Overall, my thesis aims to study multiple characteristics of *lcp1* deficient zebrafish in order to further understand the function of *lcp1* in multiple cell types that require cell motility. Using these fish, I was able to investigate three things:

1) Because L-plastin is important in immune cell function, I compared wildtype (WT) and mutant fish neutrophil abundance at two days post-fertilization. My hypothesis was that the, mutant fish would have fewer neutrophils in the caudal hematopoietic tissue (CHT), an important region of immune cell development (**Fig. 7**)

2) Next, due to the increased expression levels of *lcp1* in cancer cells, and the known correlation between this expression and the size and invasiveness of tumors, I compared the penetrance of rhabdomyosarcoma (RMS) in wildtype and *lcp1* mutants. I predicted that *lcp1* homozygous mutants would develop fewer or smaller tumors than wildtype fish (**Fig. 8**).

3) Finally, L-plastin is also highly expressed in the enveloping layer (EVL) of the zebrafish during gastrulation. It is hypothesized that this early expression of L-plastin may be correlated with the motility of the EVL cells and that *lcp1* mutants may be compensating for the

lack of *lcp1* by increasing the expression of other gene(s). I analyzed the transcriptome of both wild type and homozygous mutant fish at mid-gastrula stage to determine if any genes were differentially expressed (**Fig. 9**).

Overall, these results will further expand our knowledge of *lcp1* mutant zebrafish, as well as *lcp1's* role in hematopoiesis and cancer.

Results and Discussion

Heterozygous and homozygous mutant fish have similar numbers of neutrophils at 2 days post-fertilization.

L-plastin has been shown to be important in immune cells and is involving in the actin-bundling required for leukocyte motility. Therefore, lack of -plastin might cause poor development of such cells, or inability of these cells to migrate to the proper locations. In zebrafish, a specific leukocyte population, called neutrophils, occupies the caudal hematopoietic tissue (CHT), posterior to the yolk sac, by 2 days post-fertilization. I therefore analyzed this region to quantify neutrophils in heterozygous (+/-) and homozygous mutant (-/-) fish. Specifically, I hypothesized that homozygous mutants would have fewer neutrophils than their heterozygous siblings.

To determine if there was a difference in neutrophil number between heterozygous and homozygous mutants, clutches containing both genotypes were collected at 2 days post fertilization (dpf). Fish were fixed at this stage, and then stained using Sudan Black, which reacts with the neutrophil-specific enzyme myeloperoxidase. The result is a black precipitate in the neutrophil cell body. The fish were then bisected in order to genotype the heads (**Fig, 10A**)

and image the caudal region. Using different clutches from the same parents, I measured 21 heterozygotes and 30 mutants in this assay overall.

The tail areas were measured to determine if there was significant size variation between the two genotypes (**Fig. 10B**). Because these fish are being measured in rapidly-developing embryos, it is possible that individuals will vary in size. Clutches that contain both genotypes are the best option to use for this experiment as there will be less variation in size due to the embryos all being siblings, and the analysis can be done before determining the genotype in order to prevent bias. To measure the stained area of the tail, I used the Photoshop quick selection tool for black stained neutrophils, and then optimized manually (**Fig. 10C**). This input was measured in Image J as microns^2.

Using an unpaired t-test, there was found to be no significant difference ($p=0.668$) in the size distribution between the two genotypes (**Fig. 10C**). A similar test could be used to compare the stained area, representing neutrophils (**Fig. 10D**). This showed that there was no significant difference in the amount of neutrophils stained between the two genotypes ($p=0.895$).

This experiment shows that, for a large group of siblings reared together, homozygous mutants do not have a developmental size deficit. Rather, they achieve larval sizes comparable to heterozygotes. This result in zebrafish is similar to the mouse model of L-plastin deficiency, where all genotypes of these mice are viable (Chen et al. 2003). Additionally, when measured at 2 dpf these zebrafish mutants do not have a neutrophil deficit. This is similar to the work with the *lcp1* knockout mice which showed that there was no difference in the number of polymorphonuclear neutrophils (PMN) between wildtypes and mutants (Chen et al. 2003).

Although this experiment measures the number of neutrophils in the different genotypes, we still don't know about any other leukocyte cell type. Also, when using fixed tissue, the

motility and immune function of these neutrophils cannot be assessed. Future studies of this topic might include looking at other stages of zebrafish development, other leukocyte cell types, neutrophil response to stimuli such as wounding, or quantifying the motility of leukocytes.

Tumor penetrance between wildtype and lcp1 mutants is inconclusive

L-plastin has also been shown to be expressed in tumors, and may be related to metastasis. Therefore, lack of *lcp1* may make mutants less susceptible to the impact of an inducible tumor. Induction of rhabdomyosarcoma (RMS) in the smooth muscle of zebrafish has already been studied and the plasmids created. I utilized the RMS inducing plasmids to inject wildtype, *lcp1* heterozygous, and double mutants and compare the penetrance between the genotypes, or the number of fish that developed tumors compared to the total number of fish analyzed. Specifically, I hypothesized that the double mutant fish would have a lower penetrance of RMS when compared to wildtype or heterozygous fish.

In order to test if there is a difference in tumor penetrance between wildtype and mutant fish, I injected two plasmids, *rag2:KRAS* and *rag2:GFP,* into the fish at the one cell stage. Together, these plasmids should induce RMS in the smooth muscle tissue of the fish and the tumor cells should also be GFP+ for easy identification. After six to eight weeks of monitoring after injections, the fish were euthanized and inspected for RMS tumors, which in addition to being GFP+ should also be visible under a dissecting scope as smooth muscle tumors should appear as bumps on trunk of the fish and may distort the pigmentation cells.

The expected penetrance for wildtype fish is about 25% according to previous literature (Langenau et al. 2007). The first round of wildtype injections yielded 92 fish to analyze, and one showed signs of RMS (Fig.11 A). This is a penetrance of <1%, which is not what was expected.

Hoping to get better results with follow up injections, I re-isolated and purified the plasmids from glycerol stocks to ensure that the DNA was as clean and intact as possible. In addition to freshly isolated plasmids, the injected fish were kept in a warm incubator rather than at room temperature, where baby fish clutches are usually kept from 7-14 dpf before being placed on the running system water. This increased temperature would hopefully allow the fish to grow faster, and develop more tumors. This clutch had seventy fish to analyze, and none appeared to have developed tumors. To optimize the development of tumors, I increased the concentration of plasmids I was injecting for the final four clutches from 30 ng/mL to 50 ng/mL and also injected clutches that were 1:2:1 for WT: heterozygous: mutant genotypes in order to analyze all three genotypes at once. In total, 73 more fish were analyzed and only three appeared to have RMS. While this was an increase in penetrance when compared to the original injections, it was not adequate to determine if there was any significant difference between genotypes.

To better visualize the tumors that were found, the tumor fish, as well as some "normal" fish from the same clutch, were sectioned for histology and stained with hematoxylin and eosin before being mounted and imaged. Images were taken of the smooth muscle portion of the trunk area of the fish. (Fig. 11B). The smooth muscle cells in normal fish should appear as highly organized circles with clearly defined edges (Fig 11C), however; the tumor fish muscle cells are clearly unorganized and are running in various random patterns throughout the muscle, which is typical for muscle tumors (Fig 11D).

While optimization of the injections may need to be further studied, it is clear that the few fish that did have tumors contained muscle cells that are characteristic of RMS, and could be further studied for expression of GFP or *lcp1*. It is likely that once optimized, the lcp1 -/- zebrafish will prove a useful model for tumor development in this organism.

Possible candidates for lcp1 transcriptional adaptation in zebrafish

As previously stated in the introduction, *lcp1* mutant zebrafish are viable and reproduce even though *lcp1* is expressed in the EVL during gastrulation (**Fig 12B**). Logically, it would be hypothesized that knocking out an important gene in the gastrulation process would be lethal. Therefore, we considered that certain genes or gene families could be upregulated to compensate for the lack of *lcp1*. Strong candidates were other genes in the plastin family, other actin-binding proteins, and genes linked with *lcp1* in previous literature.

Plastin family

The plastin family contains 3 proteins: T-plastin, I-plastin, and L-plastin (LCP1) (Shinomiya 2012). Although the structure of the plastins is highly conserved, each protein has a specific function and expression pattern. While L-plastin is associated with leukocytes and cancer, T-plastin and I-plastin are associated with neurons and the intestines respectively (Shinomiya 2012). Due to the high homology between the amino acid sequences (a Clustal Omega alignment of the three human protein sequences gives an identity of 67.195% and 140 similar positions) of the different proteins, it would be logical to predict that one or both of the other plastins may be increased in expression in our mutants to compensate for the lack of LCP1.

Other cytoskeletal rearrangement genes

Other options for compensation would be genes identified in previous literature describing the role of *lcp1* in tumor cells. A study by Chaijan et al. in 2013 looked at proteins upregulated in invasive human cholangiocarcinoma cells *in vitro* when grown in basement membrane preparation (BD) Matrigel versus the same cell line not grown in BD Matrigel. In

addition to L-plastin, six other actin-binding proteins were found to be upregulated via Western blots: cytovillin 2 (ezrin), ARP3 actin-related protein 3 homolog, α-actinin-4, Adenylyl cyclase-1 associated protein, fascin, and cofilin-1 (Chaijan et al. 2014). Because these proteins can also rearrange the actin cytoskeleton and have increased expression in tumor cells similar to *lcp1*, it is possible that they could be candidates for compensation in our mutant fish. Other actin-bundling proteins that are upregulated in cancer include filamin-A, formins, mena, myosin II, T-plastin, supervillin, and villin (Stevenson et al. 2012). These are other genes to be looking for in the results of differential gene expression as they have a similar function as *lcp1* (also see **Fig. 1**).

Most of the proteins mentioned above are likely highly expressed in cancer cells due to their actin-binding qualities. Fascin, myosin, and α-actinin (*actn2* and *3* in zebrafish) are all fairly similar to LCP1 in structure, including actin binding domains, and all belong to general cytoskeletal rearrangement protein families (Winder and Ayscough 2005). Zebrafish knockouts for fascin, for example, have a severe loss of filopodia in neural crest cells, which is important for neural crest migration (Boer et al. 2015) Likewise, knockouts for *actn2* show severe skeletal muscle, cardiac, and ocular defects that cannot be rescued by *actn3* despite the similarity in sequence (Gupta et al. 2012). These proteins could also be potential targets for gene compensation as they have a similar sequence and function to LCP1

After considering candidate genes that could possibly be contributing to gene compensation in the *lcp1* mutants, RNA-seq was used to test the possible candidates or to observe if genes we had not considered were being upregulated in the mutants.

RNA-seq assumptions

It is important to note that RNA-seq runs under a set of assumptions: sequenced reads represent the cDNAs created, which represent the original levels of messenger RNAs, which

indicate protein levels, which determine the phenotype (Jacob 2014). We will assume that the assumptions hold, and will analyze the fold-changes produces by RNA-Seq as if they were valid. Previous methods for testing differential gene expression were qRT-PCR and microarray assays; however, more recent studies have shown that not only is RNA-seq as reliable as these methods, it is also better at detecting differentially expressed genes (DEGs) (Marioni et al. 2008; Wang et al. 2009; Wolff et al. 2018). For the purposes of our experiment, the WT genotype defines the "normal" expression levels of each gene and any change from this in the mutant genotype could be an indication of compensation.

Experimental design of whole-transcriptome analysis (RNA-seq)

To test for transcriptional adaptation in mutant fish, the following experiment was designed. Wild-type (WT) fish were crossed with each other to collect samples that were 100% WT and would serve as the baseline for gene expression in zebrafish. *lcp1* knockout (FX -/-) fish were also crossed with each other to collect samples that were 100% double mutants to serve as the experimental group. The results from the FX samples could then be compared to the results from the WT samples to determine possible leads for gene compensation. To include variation within a clutch, thirty embryos were homogenized to provide each sample (**Fig. 9**).

When using RNA-seq, it is also important to consider both technical and biological replicates. Technical replicates are used to test the consistency of the experimental technique, while biological replicates are used to test for variation within a biological group. My experiment includes three biological replicates for each genotype, with 30 embryos pooled in

each replicate. This design would indicate if there is extreme variation among replicates collected in this manner.

Embryonic gene expression is highly dynamic. To avoid confounding gene compensation with differences in developmental stage, it was essential to collect all embryos as stage-synchronized as possible. Zebrafish staging was based on Kimmel et al, 1995, with **Fig. 12A** serving as the reference image (Kimmel et al. 1995). All embryos were collected at 75% epiboly based on visual assessment under a dissecting scope. This stage was chosen because this is when *lcp1* is highly expressed in the outermost cells of the enveloping layer (**Fig. 12B**; (Thisse 2001; *EMBL-EMI Expression Atlas* 2018). A light micrograph of a properly-staged embryo is shown in **Fig. 12C.**

In-house RNA assays

High quality total RNA will have two distinct ribosomal bands with minimal smearing. Denaturing gel electrophoresis of the RNA samples showed consistently high quality, with no evidence of degradation (**Fig. 13A**). Based on the in-house assessment, six samples of high quality RNA (3 WT and 3 FX -/-) were sent for processing on January 8th, 2018. In the next section, I will summarize the analyses done offsite, including additional quality control, library construction, RNA sequencing, genomic alignment, and statistical analysis of differentially-expressed genes (DEGs). Additional details are provided in the Methods Section.

External vendor RNA assays

In addition to our in-house quality control, our external vendor assayed all 6 RNA samples for quality before sequencing (**Fig. 13 B &C**). The gold standard for this procedure is the Agilent 2100 Bioanalyzer (Cirera 2013). The Bioanalyzer electrophoretically separates the

components of each sample through micro channels and detects fluorescence of each component created by the gel-dye mix that is added to each sample during preparation. The results appear like bands on a traditional electrophoresis gel. These bands can be translated into peaks on a graph as well as a summary statistic called a RNA Integrity number (RIN). Higher RIN scores indicate higher quality RNA. For RNA-seq, a RIN number of 8 or above is preferred (Wang et al. 2016).

From the original six samples, a total of four were to be chosen for sequencing: two WT samples and 2 FX samples. The four samples that had the most similar RIN scores were chosen in order to keep the RNA quality between samples as consistent as possible. The four samples chosen were WT 1 and 2 and FX 1 and 2. These final four final samples, representing two biological replicate per genotype, were then processed for RNA library construction and sequencing. A general flowchart of the process is shown in **Fig. 14.** Additional details are given in the Methods section.

Per base quality

After sequencing, quality control is performed to determine the reliability of the sequences. If the individual reads are not very accurate, the data cannot be used as it would yield inaccurate global results. For each sample that was processed, we received a histogram of Phred (Phil's Read Editor) quality scores. Phred scores are calculated using the formula

$$Q = -10 \log_{10} P$$

where P = the probability of mis-calling a specific base (Technote 2011), For example, a Q of 10 has a base call accuracy of 90%, while a Q of 50 has a base call accuracy of 99.999%. A score of $Q \geq 30$ is considered "good" quality to proceed with analysis.

Figure 15 shows the histograms for our samples. The X-axis is the base pair position in the read and the Y-axis is the Phred score. The graph is divided into color sections representing the quality of the Phred score with green correlation to "good", yellow to "fair" and red to "poor". This graph shows that all of the base pair positions in all four samples are in the green "good" range. Based on these scores, we can be confident that the quality of the reads from all samples is accurate and the sequences can be aligned to the reference genome.

GC content

Another indicator of sequencing quality is GC content. The GC content of a sequence is calculated as the number of G-C bases compared to all bases (including ambiguous ones). A standard bell curve of GC content is expected for each sample. A high quality sample should have a normal distribution of GC content, with the peak being above the mean GC content of the genome in question (Babraham ; Risso et al. 2011).

Figure 16 shows the GC content for the four samples. The X-axis is the mean GC content percentage and the Y-axis is the number of reads. The blue line represents the theoretical GC content for the zebrafish genome, and the red lines are each samples actual GC content percentage. Overall, we conclude that all samples closely resemble the theoretical GC content and can be aligned confidently.

Differential gene expression: between sample similarity

Distance between samples is measured via the read counts from DESeq2 after normalization. These data can be used to analyze how similar or different samples are compared to each other (Love et al. 2014). In **Figure 17,** the darker the color, the more similar the samples, and as the color lightens, the samples are less similar. The graph shows that the two wild-type samples are each other's' closest match, as are the two FX samples. As expected, the

FX samples are more distantly related to the wild-type samples, because these are the two groups being compared. Because WT1 and WT2 were from the same WT clutch, and FX1 and FX2 were from the same FX clutch, the results are indicative of what we would expect: samples of siblings are more closely related than samples of a different line and genotype.

DESeq2

Once all the RNA fragments are sequenced and aligned, the "raw counts" are assigned to specific genes in the zebrafish genome. The next step is to analyze these data for any genes that are differentially expressed between the samples (differentially-expressed genes = DEGs). The result is shown in **Figure 18** as a heat map of all of the genes in the zebrafish genome with red indicating genes with increased expression in the wildtype and blue showing genes with increased expression in the mutant. Although this gives an overall view of the difference in expression between samples, it does not provide specific genes that could be compensating for the lack of *lcp1*.

To detect which genes have unusually altered expression, a Wald test is used to measure the statistical significance of each comparison. Fold-changes with a large Wald statistic, either positive or negative, can be considered as differentially expressed between samples. The Wald statistic then generates adjusted p-values, which returns a ranked list of the top 30 differentially expressed genes between the two genotypes (**Fig. 19**). These 30 were then clustered to identify genes that could be co-regulated. This list contains a few interesting results, including some of genes that have yet to be described (si: dkey-101k6.5, si: ch211-197g15.9, and si: ch211-133n4.4). Importantly, neither any of the plastins, nor any of the other predicted cytoskeletal

genes were in these top 30 DEGs. Normalized counts for these genes are shown in the table below:

Gene Name	WT1 count	WT2 count	FX1 count	FX2 count
T-plastin (*pls3*)	683	705	810	748
I-plastin (*pls1*)	27	24	19	24
Villin (*vil1)*	3	5	1	0
Fascin (*fscn1b*)	2	3	0	4

The next step was to further investigate some of the genes in this list of 30 for any that may be of interest for future work. To expand the scope of possible candidate genes, the entire list of differentially expressed genes was evaluated for any genes that were significantly increased in expression in mutant when compared to wildtype.

An expanded view of specific DEGs

After investigating individual differentially expressed genes, a preliminary list of twelve genes was made. These genes were chosen based on past studies, which involved cytoskeletal rearrangement and/or calcium dependence. **Figure 20** A shows the log2 fold change of these twelve genes when compared with *lcp1*. Green represents the genes that have an increased expression in mutants, and grey represents *lcp1*, which is maximal in the wildtype. The average normalized hit counts for these twelve genes and *lcp1* are represented in darker blue (**Fig, 20B)** indicating a higher hit count, whereas white indicates a very low hit count. Clearly, these genes have a large range of transcript levels, with *acin1b* being very highly expressed at 31,266

counts/sample in mutants down to *mlpha* which had only 3 counts/sample in wildtype. Bar graphs showing the normalized hit count for each sample of six selected genes are presented in **Figure 21.**

bcl2l10

*Bcl2l10 (*also called *nrz*), is part of the Bcl-2 family, which generally functions in apoptosis. However, there is some evidence that *bcl2l10* may have a function in calcium regulated cytoskeletal rearrangement (Popgeorgiev et al. 2011). Specifically, during gastrulation, *bcl2l10* mutants exhibit a premature formation of the contractile actin-myosin ring, which may be due to increased phosphorylation of MLC (myosin light chain) by way of increased calcium uptake by mitochondria, which can lead to the disorganization of the cytoskeleton (Popgeorgiev et al. 2011). Although this gene was not anticipated to be a possible candidate for transcriptional adaptation, its role in calcium-dependent cytoskeleton rearrangement corresponds with the similar function of LCP1.

mid1ip1a and midip1b

Similarly to *bcl2l10*, mutant embryos of the *mid1* genes also exhibit developmental complications (Eno et al. 2016). These embryos have F-actin organization problems during cell division. RT-PCR results indicate that *mid1ip1* is present in high levels early in development, but then is reduced in expression as development continues (Eno, Solanki, and Pelegri 2016). While this protein has not yet been studied in our mutant, perhaps the *mid1ip1* expression stays at a higher expression level further into development in order to compensate for the lack of LCP1.

plekkh1

Plekkh1, also known as *max-1*, is described as a neuronal guidance gene, but in zebrafish also functions in vascular patterning (Zhong et al. 2006). During development, vascular endothelial cells migrate to form stereotypic intersegmental blood vessels (ISV). Zhong et al. created two antisense morpholinos for *plekkh1* in zebrafish and demonstrated that coinjection of both led to complete loss of the wild-type mRNA. One day post-fertilization, injected embryos showed missing or misdirected ISVs, up to 5dpf. However, the significance of this gene in *lcp1* transcriptional adaptation remains unknown.

Verification of wildtype expression

Because expression levels can vary greatly depending on the stage of development, it is possible that staging error during collection of the samples is skewing the results. **Figure 22** shows the expression level of the twelve genes throughout development from the zygote to larval day 5. Certain genes, such as *notum1a, mid1ip1a, bcl2l10, acin1b, bnip4,* and *gpaa1* show little variation in expression around the stage collected (75% epiboly), so errors in staging would likely not affect the expression levels in these genes. However, other genes are very dynamic around 75% epiboly, and could be at a lower level in WT samples because those samples were actually collected at an earlier stage than the FX samples.

Conclusion

The three experiments described above help to further characterize the phenotype of the novel *lcp1* zebrafish line. While previous studies have looked at the immune system and cancer responses of *lcp1* mutant mice, these are some of the first experiments to analyze knockouts of

the zebrafish ortholog, and also begin to examine the transcriptome of these mutants for possible gene compensation.

Looking at the number of neutrophils between wildtype and mutants, it was hypothesized that mutants would have a fewer number of neutrophils due to the lack of *lcp1*. However, there was no significant difference found between the two genotypes. While staining with Sudan Black is relatively simple because the stage at which neutrophils are visible is known, zebrafish are transparent at that stage, and there is a specific stain for neutrophils, this experiment only investigates a single cell type of leukocytes, and is on fish that are no longer living. Future experiments that test other leukocyte cell types, such as macrophages, may produce different results. In addition, experiments that can analyze the motility and migration of leukocytes, either in general or as a response to a challenge, on live zebrafish may offer further insight into the effect of losing *lcp1* expression in the organism.

Although fish with RMS tumors were produced, the total number of fish with tumors was not high enough to form any conclusions about the difference in penetrance between the two genotypes. Inducing tumors by injecting at the one cell stage is fairly easy in zebrafish when compared to other model organisms such as mice, and RMS is an ideal solid tissue tumor as the presence of the cancer is easy to identify in the organisms without surgery. However, there are still some optimizations needed before the results that were anticipated can be produced. Further experiments could be done using a different injection apparatus, continuing to modify the concentration of plasmids injected, and keeping the growing fish at an ideal increased temperature and feeding more often. Once wildtype injections reach an ideal number of fish with tumors, mutant fish can also be injected, and the results analyzed. Based on these results,

more future experiments could be done in regards to fish with tumor's survival as well as analyzing the metastasis of the original tumors.

The generation of L-plastin mutant zebrafish that are viable and fertile allows us to investigate genes similar to *lcp1* that may not have been known to be involved in similar processes. Using RNA-seq, we have generated several leading candidates for genes that are upregulated in response to lack of *lcp1*. While this experiment itself does not provide evidence of function, it is feasible in zebrafish to further test specific genes and pathways. For example, future experiments can be done to confirm differences between WT and mutants including qRT-PCR for mRNA expression levels or Western blot analysis for protein expression levels.

Other future experiments may include creating two-locus mutants involving *lcp1* and one or more of the genes identified here. This would determine if knocking out both genes results in a lethal phenotype.

Overall, these three experiments are a starting point in the phenotypic characterization of *lcp1* mutant zebrafish and provide the foundation for future experiments regarding immune cell motility, cancer susceptibility, as well as transcriptional regulation in this organism. It is the hope that this will deliver a better understanding of *lcp1*'s function in the body and that the zebrafish can act as an educator for the role of this gene.

Methods

General zebrafish husbandry

The Institutional Animal Care and Use Committee (IACUC) of DePaul University approved all experimental protocols. Zebrafish were housed in the DePaul Research Support Facility in plastic tanks with recirculating dechlorinated tap water (pH 7-7.5, 28°C). The room

lights provide a 14-hour day: 10-hour night cycle. The adult diet was a 50:50 mix of finely ground flake food (TetraMin) and decapsulated brine shrimp eggs (American Brine Shrimp) once or twice a day. Adult fish were stocked at maximum density of 1 fish/200 mL.

Zebrafish lines

Two different zebrafish lines were used in this study. One line was the wild type ABTU strain, also called WT, and the other was the FOXTROT L-plastin mutant strain also called FX. These lines differ in that the FX (-/-) line contains a premature stop codon that was generated using the CRISPR/Cas9 gene editing system (Kell et al. 2018). When inherited in two copies, (-/-), these fish lack L-plastin protein as previously described (**Fig. 5**).

Crossing, rearing, and larval fish care

Crosses of various parental genotypes were set up for different portions of my project using the following protocol. To promote spawning, one adult male and one adult female were placed in an off-system tank overnight, and then the fertilized embryos were collected the next morning with a sterile plastic strainer (Westerfield 2000). The fertilized embryos were then incubated at 28.5°C in a Petri dish of egg water (40 g Instant Ocean/L dI water) with one drop of diluted methylene blue (1:10 in egg water) for 5-7 days. Fish to be reared were transferred to an off-system 1.4 L tank for an additional 5-7 days and fed live rotifers and/or a commercial food powder (Larval Diet, LD100; Aquatic Ecosystems). Approximately 2 weeks after fertilization, larval fish were moved to the recirculating rack and received an adult diet.

Genotyping adult fish

Adult fish were genotyped by fin-clip to confirm the parental genotype. Each fish was anaesthetized by immersion in a buffered solution of 20 mL of 0.015% Tricaine methlysulfonate

(MS-222) diluted in approximately 500mL of system water (pH 7.4) until there was no response to blunt touch (Collymore et al. 2014). The fish were then removed from the water and placed on a clean Petri dish. Using a razor blade, the posterior half of the caudal fin was excised and placed in a labeled PCR tube. The fin clips were then disintegrated for at least 2-3 hours at 55°C in 50µL lysis buffer (10 mM Tris-HCL, 50 mM KCl, 0.3% Tween-20, 0.3% Triton-X) and 1µL of Proteinase K, followed by a 10 minute incubation at 98°C to denature the enzyme. The resulting solution was stored at -20°C.

For genotyping, a primer pair specific to the CRISPR-modified region of the lcp1 gene (FWD = 5' - TGA CCT TGT CCT GCA GAT GT - 3'; REV = 5'-CTG TTT GTG CCG TCT AGT GG - 3') was used to amplify the solution of 2 µL of unpurified genomic DNA and 18 µL of PCR reaction mix (BullsEye Taq, MidSci Scientific) according to the PCR cycle described follows: 10 cycles of (92°C, 20 seconds; 68-58°C, 30 seconds; 72°C, 40 seconds), 33 cycles of (92°C, 20 seconds; 58°C, 30 seconds, 72°C, 40 seconds), and a 72°C hold for 2 minutes.

A 60 minute digestion at 55°C with 0.5 µL Bsl-1 enzyme, 1.0 µL H_2O, 1.5 µL CutSmart buffer and 12 µL of PCR product produces distinct banding patterns for wildtype, heterozygous, and mutant samples. In wild types, the diagnostic PCR product is 175 base pairs long and includes three (3) Bsl-1 restriction sites (5'–CCNNNNNN|GG–3'). A complete digestion yields fragments of 95, 56, 24, and 1 bp, respectively, however; the 1 bp band cannot be seen and the 24 bp band is often faint, so wild type alleles are best recognized as a doublet (95, 56bp). A mutant allele of LCP1 has a disrupted Bsl-1 restriction site. After digest, this yields a band of approximately 151 bp (= 95 + 56) in addition to the smaller bands generated by the normal sequence. A heterozygous sample would produce all three bands (151, 95, 56bp) and a

homozygous mutant would produce only the 151bp band. Digested fragments were resolved on a 2.5% MetaPhor agarose gel in Tris-borate-EDTA (TBE) electrophoresis buffer to observe band differences (**Fig. 10A).**

Neutrophil Staining

Crossing and collection

Crosses of heterozygous and homozygous FX mutants were set up and embryos collected the next day. The clutches from these crosses should be 50:50 heterozygous: homozygous mutant. These clutches are the best option to use for this experiment as there will be less variation in size due to the embryos all being siblings, and the analysis can be done before determining the genotype in order to prevent bias.

Fixing

After two days of soaking in PTU (1-phenyl-2-thiourea) to prevent pigment development, the embryos were fixed for 2 hours in a solution of 4% paraformaldehyde in phosphate- buffered saline (PF-PBS, pH 7.4), followed by rinsing in 1x PBS. Fixed tissue was stored at 4°C in PBS until use.

Staining with Sudan Black

To make the Sudan Black staining solution (0.03% Sudan Black in 70% ethanol), 20 mg of Sudan Black powder (Electron Microscopy Services, Ca# 21610) was mixed with 100 mL of 70% ethanol and stirred to combine. 200 uL of buffered phenol-chloroform-isoamyl alcohol (25:24:1) was added, and then the mixture was covered and stirred for 2 hours. After overnight settling, the mixture was filtered and bottled. Solution was replaced every 3-4 months.

Fixed embryos were incubated for 30 minutes in the staining solution, then excess stain was washed away with three ten-minute washed of 70% ethanol. The resulting stained embryos were then rehydrated in 3 ten-minute washes of (PBS + 0.1% Tween-20).

Bisecting

After the embryos were stained, a scalpel was used to bisect the fish between the yolk and the tail. The anterior portion of the fish was used for genotyping, while the posterior end was imaged and analyzed (Fig.7).

Genotyping

Before genotyping, the stained heads were treated in 300mM NaCl for 4 hours at 65°C. Samples were then genotyped in the same manner as previously described.

Microscopy and image analysis

Images were captured using a compound microscope and Leica imaging software. Two images were collected for each specimen using the following settings:

	Image 1:Tail Area	**Image 2: Sudan Black Area**
Condenser Position	UP	UP
Objective Lens	4x	4x
Exposure	5.22 ms	5.22 ms
Diaphragm	**0.0**	**0.5**
Gain	1.0	1.0
Saturation	1.5	1.5
Gamma	0.59	0.59

Another difference in the two images was the focus—tail area images were focused on the edge of the tail and Sudan Black area images were focused on the black stained cells of the specimen. All images included an automatic scale bar (500 microns) for area calibrations. Finally, all images were measured blindly, without knowledge of the individual genotype.

Tail area was measured in Image J (Rasband 1997-2016). After image calibration, the polygon tool was used to select an outline of the tail from the most posterior point of the yolk sac extension, perpendicularly across the body and around the end of the tail (**Fig. 10A**). The area of this selection was measured in square microns (m^2) and exported to Excel.

Images were adjusted in Photoshop to enhance the clarity of the stained cells as needed. The darkest stained pixels were manually selected, and then the selection was converted to 100% black. This image was then exported to Image J for area measurement (**Fig. 10B**). A Student's t-test was run in Prism 7 to determine if there was any statistical significance between genotypes.

Fish crossing, embryo collection and RMS injections

Wild type incrosses, FX heterozygous incrosses were used for this portion of the study (**Fig. 8).** Wild type crosses were used to test the protocol, as previous studies using the same plasmids were done on wild types and have a known penetrance (Langenau et al. 2007). Incrosses of FX +/- fish were used to produce clutches for injections that would be 1:2:1 +/+ : +/- : -/- , meaning that all possible genotypes could be observed. Fish pairs were set up the night before injections with spawning barriers in the crossing tanks in order to collect eggs laid synchronously.

Plasmids were transformed into competent E. coli cells, and then grown on plates containing ampicillin to select for cells that contained the plasmids, which contain the gene for ampicillin resistance. These colonies were then grown in LB media, and the plasmids were isolated and purified (Zymo Research MidiPrep kit, Ca#D4026). Plasmids were then linearized with XhoI at 37°C overnight as previously described by the Langenau lab (Langenau et al. 2007).

After removal of the barriers the following morning, tanks were observed every 15-20 minutes for freshly -laid eggs. Fertilized eggs were placed in a sterile agarose-filled Petri dish containing narrow trenches or slots. Up to 120 eggs were loaded onto one plate and injected at the one-cell stage with 1-2 nL of plasmid DNA solution (30-50 ng/mL each of rag2: GFP and rag2:KRAS) 1x Danieau buffer and 1% Phenol Red. All injections were into the yolk. Visual confirmation of the injection comes from the Phenol Red, which can be seen in the yolk mass, and later in the dividing blastomeres.

Fish rearing, health screening and tissue collection

Injected embryos were reared as previously described (see "Husbandry"). Fish were monitored weekly over 4-6 weeks to ensure overall health and check for tumor development. On the day before tissue collection, fish were isolated from the recirculating system and placed overnight in clean water in a 28°C incubator without feeding. This was done to clear the gut for easier visualization of tumors.

After euthanasia in ice water, a small fin clip was taken for genotyping as described previously (see 'Genotyping'). The fish were then fixed whole in 4% paraformaldehyde in phosphate- buffered saline (PF-PBS, pH 7.4) overnight at 4°C, followed by rinsing in 1x PBS. Fixed fish were then observed on a white background using a dissecting scope to survey the dorsal, ventral, and lateral sides for lumps, disruption in the parallel stripes along the body wall, or axial asymmetries, all of which can indicate a potential RMS tumor. Suspected tumors were then examined under a fluorescent dissecting scope to confirm expression of GFP (Fig. 11A). After photography, selected tissue fragments were prepared for histological examination, as described below.

Histological processing

Fixed fish that appeared to have RMS tumors were separated into head, trunk, and tail pieces using a scalpel. The tissue fragments were then placed in an acid-EDTA decalcifying solution (Cal-Ex, Fisher Scientific) overnight at 4°C. Decalcified samples were then washed in 1x PBS and dehydrated in an ethanol series up to 100% anhydrous EtOH. Samples were then washed 3x 30 minutes in amyl acetate, which serves as a bridge between the ethanol and the molten paraffin used for embedding.

Specimens were passed through 3 changes of melted paraffin (58-6°C for 30 minutes each), then placed in warm vacuum cabinet to remove excess air. Molds were sprayed with mold release, and then a small amount of wax was added to the bottom of the molds and cooled until sticky. Tissue fragments could then be placed into mold and oriented in the desired manner before filling the mold with wax and allowing it to set overnight. Embedded samples were unmolded the next day and stored in a dust free container at room temperature until sectioning.

Paraffin blocks were sectioned at 10 microns using a manual rotary microtome (American Optical 820). Cut sections were floated on warm distilled water, then mounted on coated glass slides (Colorfrost Plus, Fisher Scientific). Slides were then dried overnight at room temperature under a dust cover. Hematoxylin and eosin staining was applied using the following protocol: 2 rounds of Citrasolv (Fisher brand, Ca# 04-355-121), 1 minute; series of 100%, 95%, 70%, and 50% ethanol for 1 minute; dI water, 1 minute; hematoxylin (Electron Microscopy Sciences, Ca# 26030-20), 1 minute; running tap water until solution runs clear; acid alcohol, <10

seconds; dI water, 1 minute; eosin (Electron Microscopy Sciences, Ca# 26762-01), <10 seconds; 95% and then 100% ethanol, 1 minute; 2 rounds of Citrasolv, 1 minute.

Once stained, coverslips were mounted onto slides using DPX mounting medium (Sigma Aldrich) and set flat overnight to dry.

Embryo collection and genotyping for RNA-seq

Genotyped males and females were paired to produce 100% wild type or 100% homozygous mutant embryos (FX -/-), respectively. Embryos were incubated in a Petri dish at 28°C for 5-6 hours until 70 epiboly. A single RNA sample was obtained by placing 30 embryos from the same clutch into a 1.5mL Eppendorf tube for collection. 300µL of TRI-zol reagent was added and embryos were then homogenized for 10-15 seconds using a motorized micropestle. Homogenates were stored at -80°C until RNA extraction.

To confirm the genotype of each clutch (100% WT or 100% FX -/-), eight embryos per clutch were also collected in individual PCR tubes and genotyped as previously described (see 'genotyping'). Once the genotype of the samples was confirmed, RNA could be isolated and purified.

RNA isolation and purification

Total RNA was extracted and purified using the Direct-zol RNA MiniPrep kit (Zymo Research R2070) following the manufacturer's instructions. Briefly, the RNA samples were thawed on ice, an equal volume of 100% ethanol was added, and the samples were then vortexed and spun down. The mixture was then transferred to a spin column in a collection tube and centrifuged at 14,000 rpm for 1 minute with flow through discarded. Columns were then washed with RNA Wash Buffer and centrifuged again, then treated with a master mix of (5µL DNase I +

75µL of DNA digestion buffer) at room temperature for 15 minutes. This step ensures that no genomic DNA contaminates the RNA samples. The columns were then washed with RNA PreWash, RNA Wash Buffer, and finally eluted in 50µL of nuclease-free water. Purified RNA was stored at -80°C until samples were tested for quality control.

RNA quality control

In house assays

RNA concentration and purity were confirmed using two in-house methods: spectrophotometry and agarose gel electrophoresis. Sample concentration was assessed using a Nanodrop 2000c to measure the 260 nm /280 nm absorbance ratio. Purity was assessed using the 260 nm /230 nm absorbance ratio to confirm that the sample was free from contamination from extraction reagents or other contaminants. If samples had a concentration and purity ratios that exceeded the requirements of the sequencing vendor (concentration >50 ng/ µL, A260/280=1.8-2.2), they were further examined for RNA quality on a denaturing agarose gel.

Denaturing gel analysis of total RNA

A small portion of each RNA sample (approximately 300ng of RNA) was added to RNA loading dye (New England Biolabs, #B0363A) and heated at 70°C for 10 minutes to denature. The denatured RNA was placed on ice for 1-3 minutes, then run on a fresh 1% lithium boride agarose gel with a DNA size marker (Hi-Lo Marker, Bionexus). Electrophoresis was at 220V for 30 minutes.

External vendor library preparation and sequencing

Total RNA was reverse-transcribed to cDNA for stability, and fragmented to 200-500 bp using either DNase I treatment or sonication. Next, the RNA fragments were tagged with adapters (Wang, Gerstein, and Snyder 2009). Adapters are short oligonucleotides used as sample-specific "barcodes". The cDNA fragments in each sample are given a specific adapter sequence to identify the reads specific to that sample. This allows for all of the samples to be run in one sequencing lane, and eliminates the need for technical replicates because all samples are being run on the same machine, in the same lane, and at the same time (Schiemer 2011).

The fragments are then sequenced using an Illumina sequencer. The Illumina sequencer uses fluorescent bases to determine the sequence of each fragment via complementation (Quail et al. 2012). Typically, 30 million reads is needed to capture most of the expressed genes in sample for differential gene expression analysis (Korpelainen et al. 2015). If all of the samples have over 30 million sequenced reads, this means that the samples have a high enough read number to move forward with bioinformatics analysis. Reads were as follows: WT1 and 2 = >34 million, FX 1 and 2 = >32 million.

Aligning sequenced reads to the genome.

After RNA is sequenced, fragments were aligned to the zebrafish reference genome using the program Spliced Transcripts Alignment to a Reference (STAR) (Dobin et al. 2013). This program analyzes each fragment, or read, to assign it a specific genomic location. STAR is unique compared to other alignment programs because it will find the longest portion of a fragment that matches the genome, and then analyze if the rest of the fragment matches a portion

of the genome close by. For example, the first part of the fragment may correspond to exon 3 of a gene, and the rest of the fragment corresponds to exon 4. STAR ensures that the portion corresponding to exon 4 is analyzed first in order to prevent mismatches to other unrelated portions of the genome. This then results in data for the number of "hits" for the 5' and 3' untranslated region (UTR), and various exons for each gene.

Differential gene expression

The bioinformatics of differential gene expression was performed with the program, DESeq2, which uses the negative binomial test on variance estimated (Love, Huber, and Anders 2014). DESeq2 can be run in R/BioConductor and the package is freely available (Love, Huber, and Anders 2014).

DESeq2 uses a count matrix, K, with one row for each gene, i, and one column for each sample, j. Read counts K_{ij} can be modeled using a negative binomial distribution with mean, μ_{ij}, and dispersion, $\alpha_{i.}$. This process essentially takes the raw count data from STAR and normalizes it into read counts that can be used to analyze similarities or differences in gene expression between samples (Love, Huber, and Anders 2014).

References

Amos, L. A. and Amos, W. B. 1991. *Molecules of the Cytoskeleton*. Edited by C. J. Skidmore, *Guilford Series*. New York: Guilford Press.

Arpin, M., Friederich, E., Algrain, M., Vernel, F., and Louvard, D. 1994. Functional Differences Between L-plastin and T-plastin Isoforms. *Journal of Cell Biology* 127 (6):1995-2008.

Babraham, B. Per Sequence GC Content. babraham.ac.uk.

Balciunas, D. 2018. Fish mutant, where is thy phenotype? *PLoS Genet* 14 (2):e1007197.

Basten, S. G., Davis, E. E., Gillis, A. J., van Rooijen, E., Stoop, H., Babala, N., Logister, I., Heath, Z. G., Jonges, T. N., Katsanis, N., Voest, E. E., van Eeden, F. J., Medema, R. H., Ketting, R. F., Schulte-Merker, S., Looijenga, L. H., and Giles, R. H. 2013. Mutations in LRRC50 predispose zebrafish and humans to seminomas. *PLoS Genet* 9 (4):e1003384.

Boer, E. F., Howell, E. D., Schilling, T. F., Jette, C. A., and Stewart, R. A. 2015. Fascin1-dependent Filopodia are required for directional migration of a subset of neural crest cells. *PLoS Genet* 11 (1):e1004946.

Bosseler, M., Marani, V., Broukou, A., Lequeux, A., Kaoma, T., Schlesser, V., François, J. H., Palissot, V., Berchem, G. J., Aouali, N., and Janji, B. 2018. Inhibition of HIF1α-Dependent Upregulation of Phospho-l-Plastin Resensitizes Multiple Myeloma Cells to Frontline Therapy. *Int J Mol Sci* 19 (6).

Boxer, L. A., Richardson, S., and Floyd, A. 1976. Identification of actin-binding protein in membrane of polymorphonuclear leukocytes. *Nature* 263 (5574):249-51.

Bray, D. 1992. *Cell Movements*. New York & London: Garland Publishing, Inc.

Chaijan, S., Roytrakul, S., Mutirangura, A., and Leelawat, K. 2014. Matrigel induces L-plastin expression and promotes L-plastin-dependent invasion in human cholangiocarcinoma cells. *Oncol Lett* 8 (3):993-1000.

Chen, E. Y. and Langenau, D. M. 2011. Zebrafish models of rhabdomyosarcoma. *Methods Cell Biol* 105:383-402.

Chen, H., Mocsai, A., Zhang, H., Ding, R. X., Morisaki, J. H., White, M., Rothfork, J. M., Heiser, P., Colucci-Guyon, E., Lowell, C. A., Gresham, H. D., Allen, P. M., and Brown, E. J. 2003. Role for plastin in host defense distinguishes integrin signaling from cell adhesion and spreading. *Immunity* 19 (1):95-104.

Cirera, S. 2013. Highly efficient method for isolation of total RNA from adipose tissue. *BMC Res Notes* 6:472.

Clarkson, E., Costa, C. F., and Machesky, L. M. 2004. Congenital myopathies: diseases of the actin cytoskeleton. *J Pathol* 204 (4):407-17.

Collymore, C., Tolwani, A., Lieggi, C., and Rasmussen, S. 2014. Efficacy and safety of 5 anesthetics in adult zebrafish (Danio rerio). *J Am Assoc Lab Anim Sci* 53 (2):198-203.

Cooper, J. A., Buhle, E. L., Walker, S. B., Tsong, T. Y., and Pollard, T. D. 1983. Kinetic evidence for a monomer activation step in actin polymerization. *Biochemistry* 22 (9):2193-202.

Crowhurst, M. O., Layton, J. E., and Lieschke, G. J. 2002. Developmental biology of zebrafish myeloid cells. *Int J Dev Biol* 46 (4):483-92.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29 (1):15-21.

Dor-On, E., Raviv, S., Cohen, Y., Adir, O., Padmanabhan, K., and Luxenburg, C. 2017. T-plastin is essential for basement membrane assembly and epidermal morphogenesis. *Sci Signal* 10 (481).

Elzinga, M., Collins, J. H., Kuehl, W. M., and Adelstein, R. S. 1973. Complete amino-acid sequence of actin of rabbit skeletal muscle. *Proc Natl Acad Sci U S A* 70 (9):2687-91.

EMBL-EMI Expression Atlas. 2018. 2018 [cited 2018]. Available from https://www.ebi.ac.uk/gxa/genes/ensdarg00000023188?bs=%7B%22danio rerio%22%3A%5B%22DEVELOPMENTAL_STAGE%22%5D%7D&ds=%7B%22kingdom%22%3A%5B%22animals%22%5D%7D - baseline.

Eno, C., Solanki, B., and Pelegri, F. 2016. aura (mid1ip1l) regulates the cytoskeleton at the zebrafish egg-to-embryo transition. *Development* 143 (9):1585-99.

Fletcher, D. A. and Mullins, R. D. 2010. Cell mechanics and the cytoskeleton. *Nature* 463 (7280):485-92.

Galkin, V. E., Orlova, A., Cherepanova, O., Lebart, M. C., and Egelman, E. H. 2008. High-resolution cryo-EM structure of the F-actin-fimbrin/plastin ABD2 complex. *Proc Natl Acad Sci U S A* 105 (5):1494-8.

Goldstein, D., Djeu, J., Latter, G., Burbeck, S., and Leavitt, J. 1985. Abundant synthesis of the transformation-induced protein of neoplastic human fibroblasts, plastin, in normal lymphocytes. *Cancer Res* 45 (11 Pt 2):5643-7.

Grether, G. F. 2005. Environmental change, phenotypic plasticity, and genetic compensation. *Am Nat* 166 (4):E115-23.

Gupta, V., Discenza, M., Guyon, J. R., Kunkel, L. M., and Beggs, A. H. 2012. α-Actinin-2 deficiency results in sarcomeric defects in zebrafish that cannot be rescued by α-actinin-3 revealing functional differences between sarcomeric isoforms. *FASEB J* 26 (5):1892-908.

Hashimoto, S., Nagai, S., Sese, J., Suzuki, T., Obata, A., Sato, T., Toyoda, N., Dong, H. Y., Kurachi, M., Nagahata, T., Shizuno, K., Morishita, S., and Matsushima, K. 2003. Gene expression profile in human leukocytes. *Blood* 101 (9):3509-13.

Herbomel, P., Thisse, B., and Thisse, C. 1999. Ontogeny and behaviour of early macrophages in the zebrafish embryo. *Development* 126 (17):3735-45.

Huiting, L. N., Laroche, F., and Feng, H. 2015. The Zebrafish as a Tool to Cancer Drug Discovery. *Austin J Pharmacol Ther* 3 (2):1069.

Jacob, J. 2014. RNA-seq: general concept, goal, and experimental design. https://www.slideshare.net/joachimjacob/1rna-seqpart1working-tothegoal: BITS.

Janji, B., Giganti, A., De Corte, V., Catillon, M., Bruyneel, E., Lentz, D., Plastino, J., Gettemans, J., and Friederich, E. 2006. Phosphorylation on Ser5 increases the F-actin-binding activity of L-plastin and promotes its targeting to sites of actin assembly in cells. *J Cell Sci* 119 (Pt 9):1947-60.

Kang, S., Shim, H. S., Lee, J. S., Kim, D. S., Kim, H. Y., Hong, S. H., Kim, P. S., Yoon, J. H., and Cho, N. H. 2010. Molecular Proteomics Imaging of Tumor Interfaces by Mass Spectrometry. *Journal of Proteome Research* 9 (2):1157-1164.

Kell, M. J., Riccio, R. E., Baumgartner, E. A., Compton, Z. J., Pecorin, P. J., Mitchell, T. A., Topczewski, J., and LeClair, E. E. 2018. Targeted deletion of the zebrafish actin-bundling protein L-plastin (lcp1). *PLoS One* 13 (1):e0190353.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn* 203 (3):253-310.

Klemke, M., Rafael, M. T., Wabnitz, G. H., Weschenfelder, T., Konstandin, M. H., Garbi, N., Autschbach, F., Hartschuh, W., and Samstag, Y. 2007. Phosphorylation of ectopically expressed L-plastin enhances invasiveness of human melanoma cells. *International Journal of Cancer* 120 (12):2590-2599.

Koide, N., Kasamatsu, A., Endo-Sakamoto, Y., Ishida, S., Shimizu, T., Kimura, Y., Miyamoto, I., Yoshimura, S., Shiiba, M., Tanzawa, H., and Uzawa, K. 2017. Evidence for Critical Role of Lymphocyte Cytosolic Protein 1 in Oral Cancer. *Sci Rep* 7:43379.

Korpelainen, E., Tuimala, J., Somervuo, P., Huss, M., and Wong, G. 2015. *RNA-seq Data Analysis- A Practical Approach*. Edited by N. F. Britton, X. Lin, H. M. Safer, M. V. Schneider, M. Singh and A. Tramontano, *Mathematical and Computational Biology Series*. Boca Raton, FL: CRC Press

Langenau, D. M., Keefe, M. D., Storer, N. Y., Guyon, J. R., Kutok, J. L., Le, X., Goessling, W., Neuberg, D. S., Kunkel, L. M., and Zon, L. I. 2007. Effects of RAS on the genesis of embryonal rhabdomyosarcoma. *Genes Dev* 21 (11):1382-95.

Le Guyader, D., Redd, M. J., Colucci-Guyon, E., Murayama, E., Kissa, K., Briolat, V., Mordelet, E., Zapata, A., Shinomiya, H., and Herbomel, P. 2008. Origins and unconventional behavior of neutrophils in developing zebrafish. *Blood* 111 (1):132-41.

Lieschke, G. J. and Currie, P. D. 2007. Animal models of human disease: zebrafish swim into view. *Nat Rev Genet* 8 (5):353-67.

Lin, C. S., Aebersold, R. H., Kent, S. B., Varma, M., and Leavitt, J. 1988. Molecular cloning and characterization of plastin, a human leukocyte protein expressed in transformed human fibroblasts. *Mol Cell Biol* 8 (11):4659-68.

Lin, C. S., Chen, Z. P., Park, T., Ghosh, K., and Leavitt, J. 1993. Characterization of the human L-plastin gene promoter in normal and neoplastic cells. *Journal of Biological Chemistry* 268 (4):2793-2801.

Lin, C. S., Park, T., Chen, Z. P., and Leavitt, J. 1993. Human plastin genes- comparative gene structure, chromosome location, and differential expression in normal and neoplastic cells.. *Journal of Biological Chemistry* 268 (4):2781-2792.

Lin, C. S. 1993. Human plastin genes. Comparative gene structure, chromosome location, and differential expression in normal and neoplastic cells. *J Biol Chem* 268 (4):2781-92.

Lin, C. S., Shen, W., Chen, Z. P., Tu, Y. H., and Matsudaira, P. 1994. Identification of I-plastin, a human fimbrin isoform expressed in intestine and kidney. *Mol Cell Biol* 14 (4):2457-67.

Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15 (12):550.

Ma, T., Sadashivaiah, K., Madayiputhiya, N., and Chellaiah, M. 2010. Regulation of sealing ring formation by L-plastin and cortactin in osteoclasts. *J Biol Chem* 285 (39):29911-24.

Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M., and Gilad, Y. 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 18 (9):1509-17.

Mathias, J. R., Dodd, M. E., Walters, K. B., Yoo, S. K., Ranheim, E. A., and Huttenlocher, A. 2009. Characterization of zebrafish larval inflammatory macrophages. *Dev Comp Immunol* 33 (11):1212-7.

Mione, M. C. and Trede, N. S. 2010. The zebrafish as a model for cancer. *Dis Model Mech* 3 (9-10):517-23.

Mirbahai, L., Williams, T. D., Zhan, H., Gong, Z., and Chipman, J. K. 2011. Comprehensive profiling of zebrafish hepatic proximal promoter CpG island methylation and its modification during chemical carcinogenesis. *BMC Genomics* 12:3.

Mizgirev, I. and Revskoy, S. 2010. Generation of clonal zebrafish lines and transplantable hepatic tumors. *Nat Protoc* 5 (3):383-94.

Morley, S. C. 2013. The actin-bundling protein L-plastin supports T-cell motility and activation. *Immunol Rev* 256 (1):48-62.

Otsuka, M., Kato, M., Yoshikawa, T., Chen, H., Brown, E. J., Masuho, Y., Omata, M., and Seki, N. 2001. Differential expression of the L-plastin gene in human colorectal cancer progression and metastasis. *Biochemical and Biophysical Research Communications* 289 (4):876-881.

Otto, J. J. 1994. Actin-bundling proteins. *Curr Opin Cell Biol* 6 (1):105-9.

Park, T. S., Chen, Z. P., and Leavitt, J. 1994. Activation of the leukocyte plastin gene occurs in most human cancer cells. *Cancer Research* 54 (7):1775-1781.

Paul, A. S. and Pollard, T. D. 2009. Review of the mechanism of processive actin filament elongation by formins. *Cell Motil Cytoskeleton* 66 (8):606-17.

Peng, X. Y., Won, J. H., Rutherford, T., Fujii, T., Zelterman, D., Pizzorno, G., Sapi, E., Leavitt, J., Kacinski, B., Crystal, R., Schwartz, P., and Deisseroth, A. 2001. The use of the L-plastin promoter for adenoviral-mediated, tumor-specific gene expression in ovarian and bladder cancer cell lines. *Cancer Research* 61 (11):4405-4413.

Popgeorgiev, N., Bonneau, B., Ferri, K. F., Prudent, J., Thibaut, J., and Gillet, G. 2011. The apoptotic regulator Nrz controls cytoskeletal dynamics via the regulation of Ca2+ trafficking in the zebrafish blastula. *Dev Cell* 20 (5):663-76.

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., Bertoni, A., Swerdlow, H. P., and Gu, Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341.

Rasband, W. *ImageJ.* http://imagej.nih.gov/ij/. U.S. National Institutes of Health 1997-2016 [cited. Available from http://imagej.nih.gov/ij/.

Riplinger, S. M., Wabnitz, G. H., Kirchgessner, H., Jahraus, B., Lasitschka, F., Schulte, B., van der Pluijm, G., van der Horst, G., Hammerling, G. J., Nakchbandi, I., and Samstag, Y. 2014. Metastasis of prostate cancer and melanoma cells in a preclinical in vivo mouse model is enhanced by L-plastin expression and phosphorylation. *Mol Cancer* 13:10.

Risso, D., Schwartz, K., Sherlock, G., and Dudoit, S. 2011. GC-content normalization for RNA-Seq data. *BMC Bioinformatics* 12:480.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., and Stainier, D. Y. 2015. Genetic compensation induced by deleterious mutations but not gene knockdowns. *Nature* 524 (7564):230-3.

Schiemer, J. 2011. Illumina TruSeq DNA Adapters De-Mystified. 2011 Tufts University Core Facility.

Schliwa, M. 1986. *The Cytoskeleton*. Edited by M. Alfert, W. Beermann, L. Goldstein and K. R. Porter. 13 vols. Vol. 13. Austria: Adolf Holzhausens Nfg., Wien.

Schwebach, C. L., Agrawal, R., Lindert, S., Kudryashova, E., and Kudryashov, D. S. 2017. The Roles of Actin-Binding Domains 1 and 2 in the Calcium-Dependent Regulation of Actin Filament Bundling by Human Plastins. *J Mol Biol* 429 (16):2490-2508.

Shinomiya, H. 2012. Plastin family of actin-bundling proteins: its functions in leukocytes, neurons, intestines, and cancer. *Int J Cell Biol* 2012:213492.
Stevenson, R. P., Veltman, D., and Machesky, L. M. 2012. Actin-bundling proteins in cancer progression at a glance. *J Cell Sci* 125 (Pt 5):1073-9.
Szent-Györgyi, A. G. 2004. The early history of the biochemistry of muscle contraction. *J Gen Physiol* 123 (6):631-41.
Technote, I. c. 2018. *Quality score for next-generation sequencing* 2011 [cited 2018]. Available from https://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf.
Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A., Steffan, T., Charbonnier, X.Q. and Thisse, C. 2001. Expression of the zebrafish genome during embryogenesis (NIH R01 RR15402) . ZFIN Direct Data Submission. .
Todd, E. M., Deady, L. E., and Morley, S. C. 2013. Intrinsic T- and B-cell defects impair T-cell-dependent antibody responses in mice lacking the actin-bundling protein L-plastin. *Eur J Immunol* 43 (7):1735-44.
Todd, E. M., Zhou, J. Y., Szasz, T. P., Deady, L. E., D'Angelo, J. A., Cheung, M. D., Kim, A. H., and Morley, S. C. 2016. Alveolar macrophage development in mice requires L-plastin for cellular localization and retention within alveoli. *Blood*.
Traver, D., Herbomel, P., Patton, E. E., Murphey, R. D., Yoder, J. A., Litman, G. W., Catic, A., Amemiya, C. T., Zon, L. I., and Trede, N. S. 2003. The zebrafish as a model organism to study development of the immune system. *Adv Immunol* 81:253-330.
Trede, N. S., Langenau, D. M., Traver, D., Look, A. T., and Zon, L. I. 2004. The use of zebrafish to understand immunity. *Immunity* 20 (4):367-79.
Wang, C., Morley, S. C., Donermeyer, D., Peng, I., Lee, W. P., Devoss, J., Danilenko, D. M., Lin, Z., Zhang, J., Zhou, J., Allen, P. M., and Brown, E. J. 2010. Actin-bundling protein L-plastin regulates T cell activation. *J Immunol* 185 (12):7487-97.
Wang, L., Nie, J., Sicotte, H., Li, Y., Eckel-Passow, J. E., Dasari, S., Vedell, P. T., Barman, P., Weinshiboum, R., Jen, J., Huang, H., Kohli, M., and Kocher, J. P. 2016. Measure transcript integrity using RNA-seq data. *BMC Bioinformatics* 17:58.
Wang, Z., Gerstein, M., and Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10 (1):57-63.
Westerfield, M. 2000. Ch. 1 General Methods for Zebrafish Care. In *The zebrafish book: a guide for the laboratory use of zebrafish (Danio rerio)*. Eugene, OR: Univ. of Oregon Press.
Winder, S. J. and Ayscough, K. R. 2005. Actin-binding proteins. *J Cell Sci* 118 (Pt 4):651-4.
Wolff, A., Bayerlová, M., Gaedcke, J., Kube, D., and Beißbarth, T. 2018. A comparative study of RNA-Seq and microarray data analysis on the two examples of rectal-cancer patients and Burkitt Lymphoma cells. *PLoS One* 13 (5):e0197162.
Yen, J., White, R. M., and Stemple, D. L. 2014. Zebrafish models of cancer: progress and future challenges. *Curr Opin Genet Dev* 24:38-45.
Zheng, J. P., Rudra-Ganguly, N., Powell, W. C., and Roy-Burman, P. 1999. Suppression of prostate carcinoma cell invasion by expression of antisense L-plastin gene. *American Journal of Pathology* 155 (1):115-122.
Zhong, H., Wu, X., Huang, H., Fan, Q., Zhu, Z., and Lin, S. 2006. Vertebrate MAX-1 is required for vascular patterning in zebrafish. *Proc Natl Acad Sci U S A* 103 (45):16800-5.
Zhou, J. Y., Szasz, T. P., Stewart-Hutchinson, P. J., Sivapalan, J., Todd, E. M., Deady, L. E., Cooper, J. A., Onken, M. D., and Morley, S. C. 2016. L-Plastin promotes podosome longevity and supports macrophage motility. *Mol Immunol* 78:79-88.

Figures

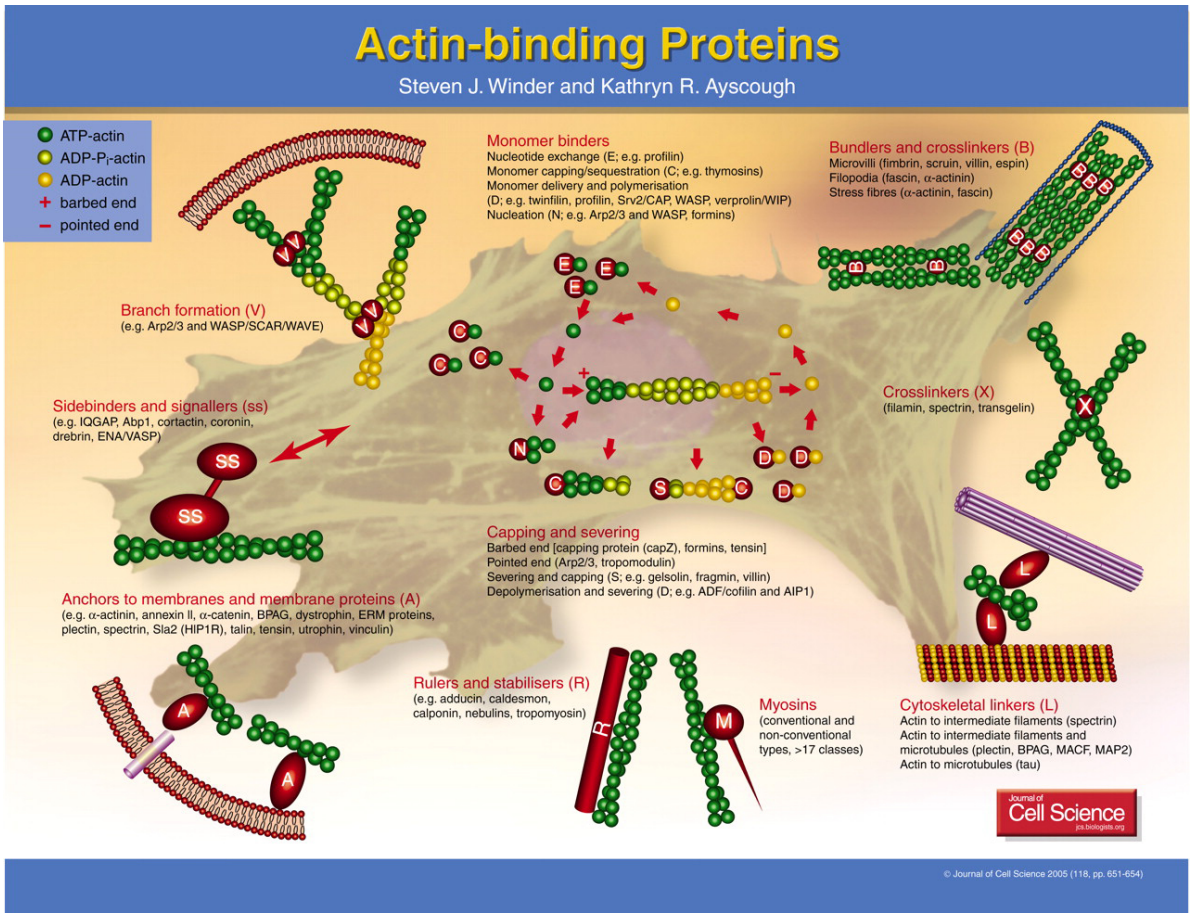

Figure 1:Actin-binding Proteins

The figure shows various proteins that can bind actin for specific functions.
Image Source: Actin-binding proteins, Steven J. Winder, Kathryn R. Ayscough, Journal of Cell Science 2005 118: 651-654; doi: 10.1242/jcs.01670

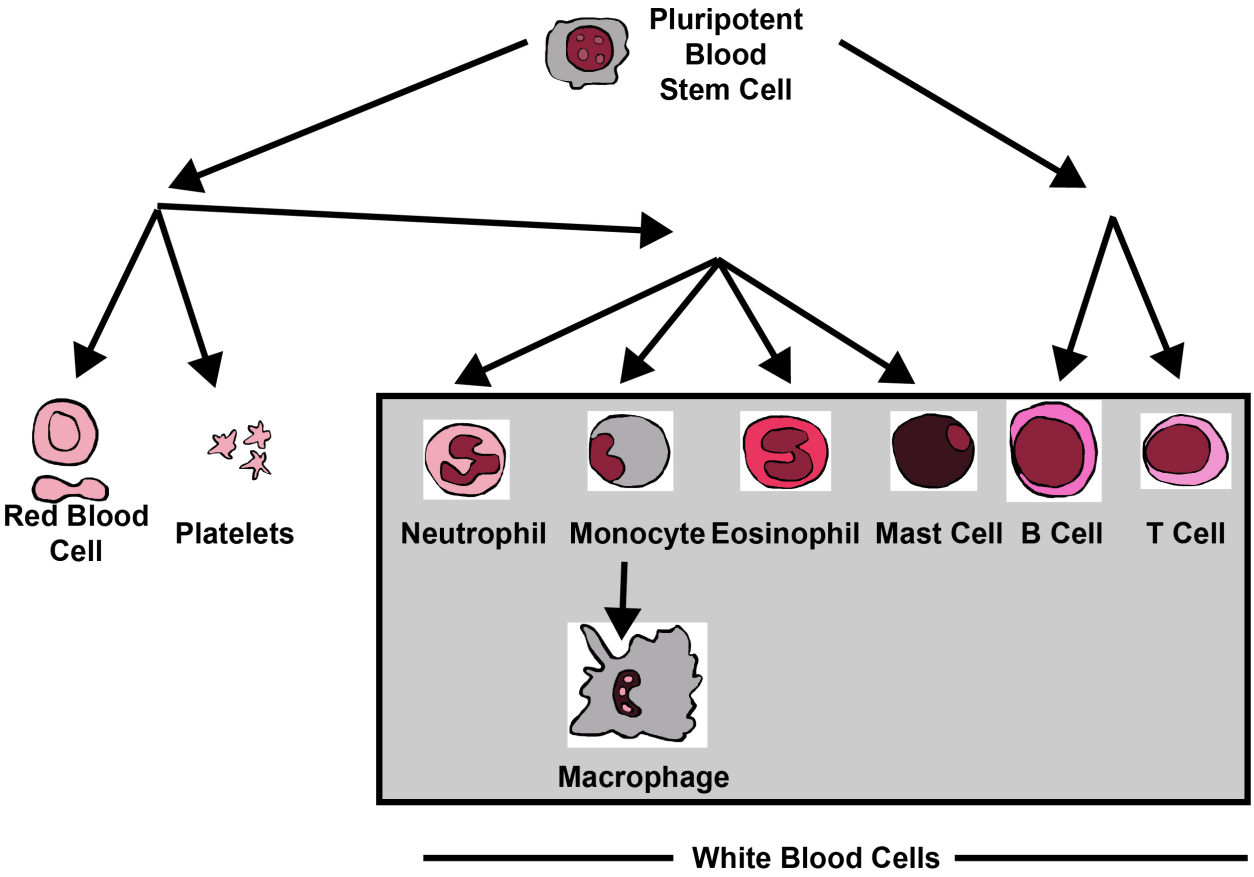

Figure 2: Lineage hierarchy of blood and immune cells.

White blood cells, or leukocytes, all express *lcp1* and are derived from pluripotent blood stem cells. However, other cell types derived from these same stem cells, such as red blood cells and platelets, do not express *lcp1*. Image Source: (modified from https://en.wikipedia.org/wiki/User:A._Rad.

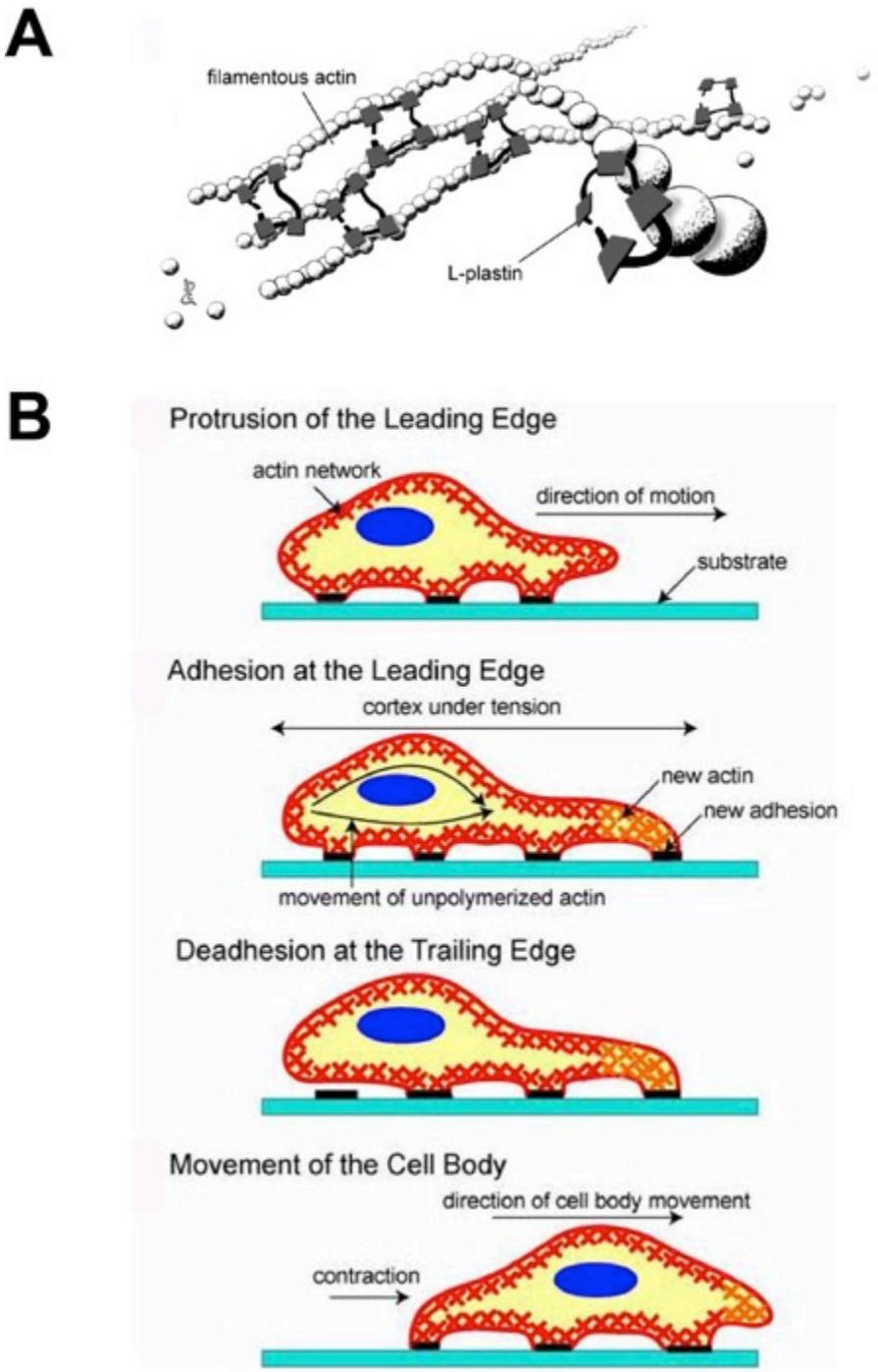

Figure 3: Structure and function of L-plastin, a cytoplasmic actin-bundling protein.

A) L-plastin binds to filamentous actin and creates bundles within the cytoplasm. Image Source: (Kell et al., 2018). B) Rearrangement of the cytoskeleton allows for the motility of cells via lamellipodia and filopodia. Image Source: (Ananthakrishnan and Ehrlicher, 2007).

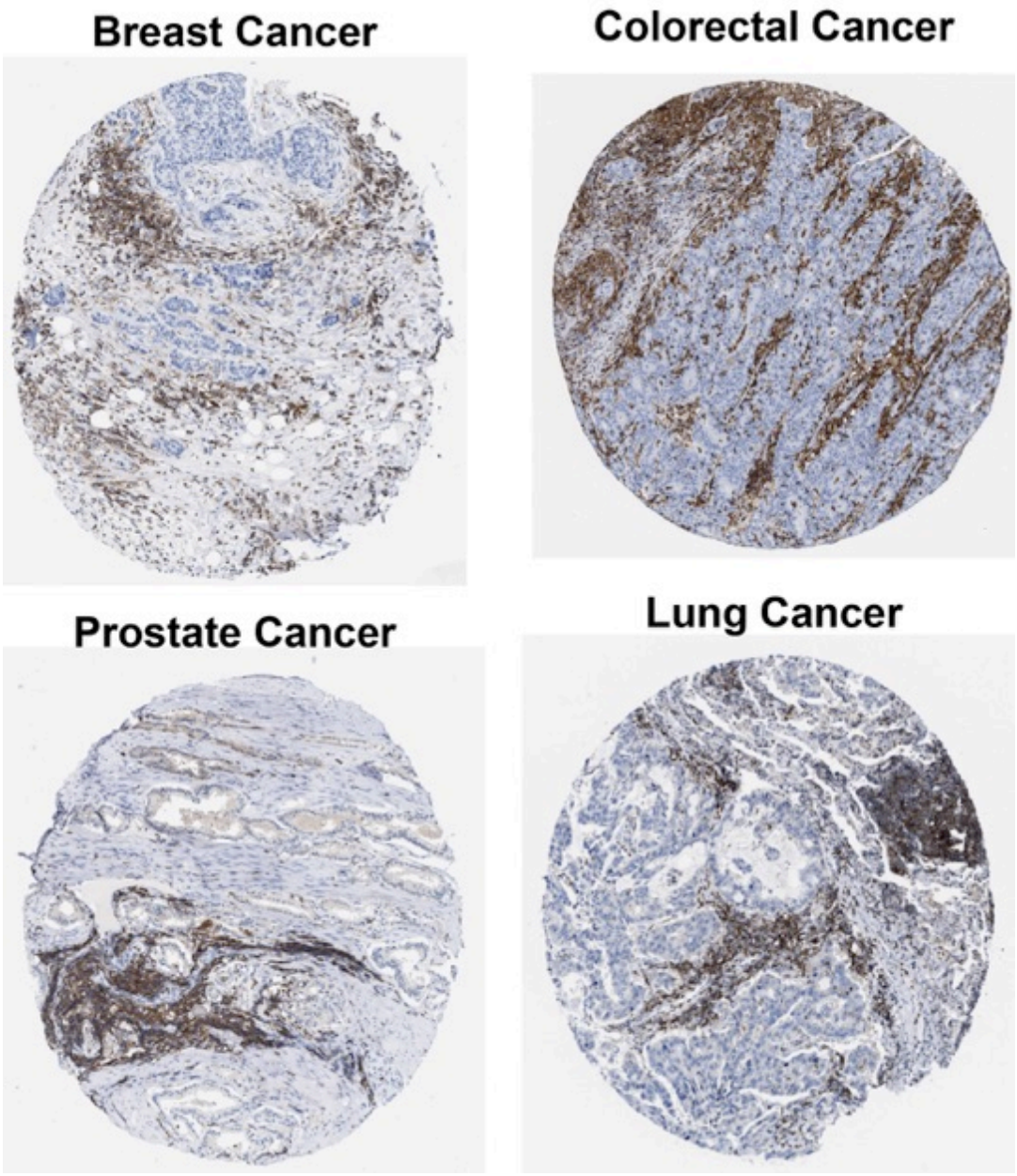

Figure 4: Immunohistochemical staining for L-plastin protein in four human tumors.

Sections were stained for LCP1 using antibody HPA019493 or CAB020673, followed by secondaries and DAB substrate (give a bit more detail here). Brown cells = L-plastin (+) cells; blue = hematoxylin-stained nuclei. Image Source: (The Human Protein Atlas).

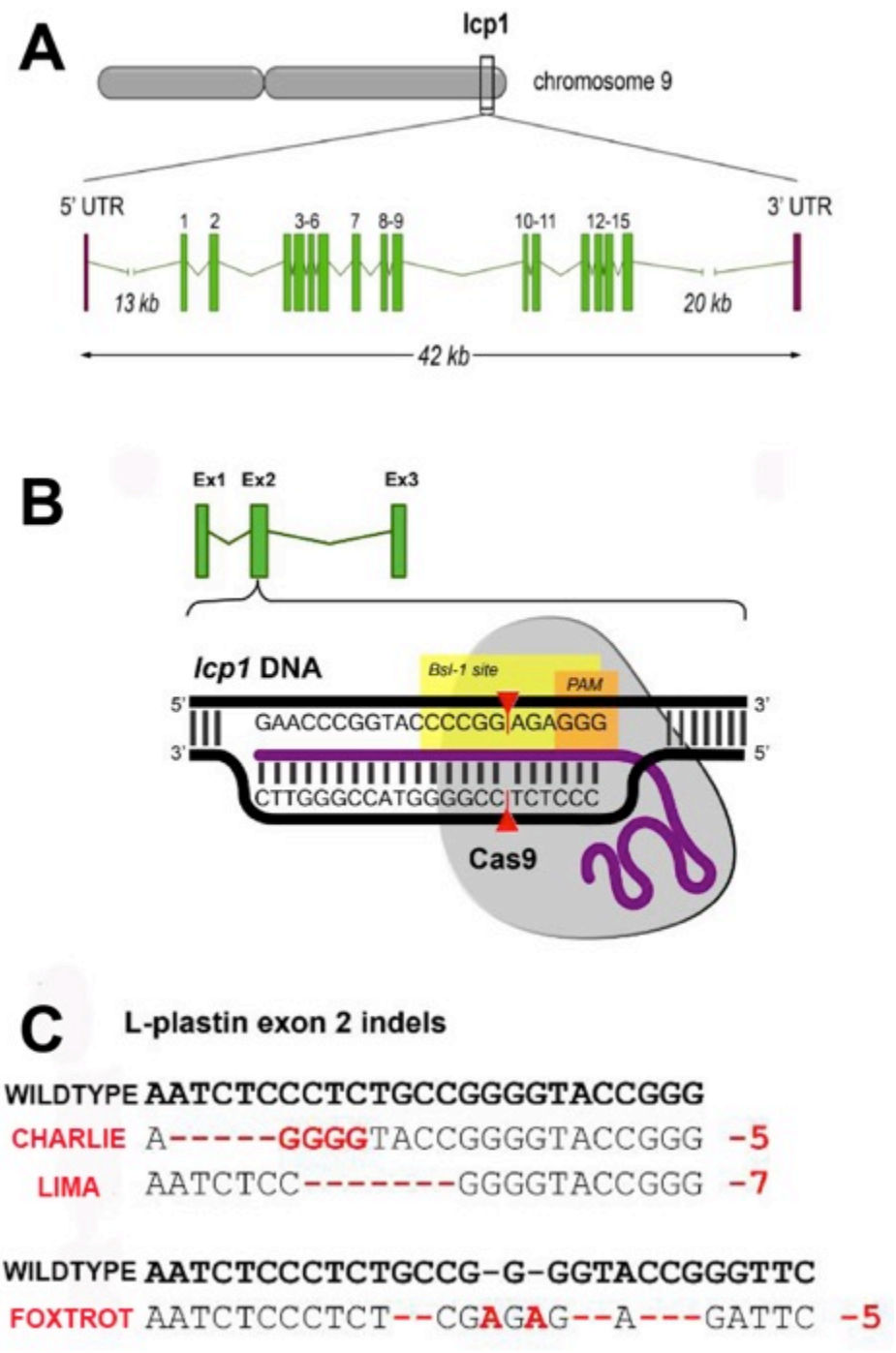

Figure 5: Genetic engineering of *lcp1* mutant zebrafish.
A) Location of L-plastin on zebrafish chromosome 9 and exon-intron structure of the locus. The most 5' and 3'exons are untranslated (5' and 3' UTR) flank the coding exons (#1-15). B) Schematic of the CRISPR/Cas9 target site within exon 2. The exon 2 guide RNA (purple) binds the complementary genomic DNA (black) and docks with the Cas9 nuclease (gray). Upstream of the protospacer-adjacent motif (PAM, in orange), the Cas9 cleavage site targets a Bsl1restriction site (yellow), causing indels. C) Alignment of wild type and edited alleles of zebrafish *lcp1*. CH= 'Charlie', a 5 bp deletion; LM= 'Lima", a 7-bp deletion; FX= 'Foxtrot', a net 5 bp deletion. Image Source: (Kell et al., 2018).

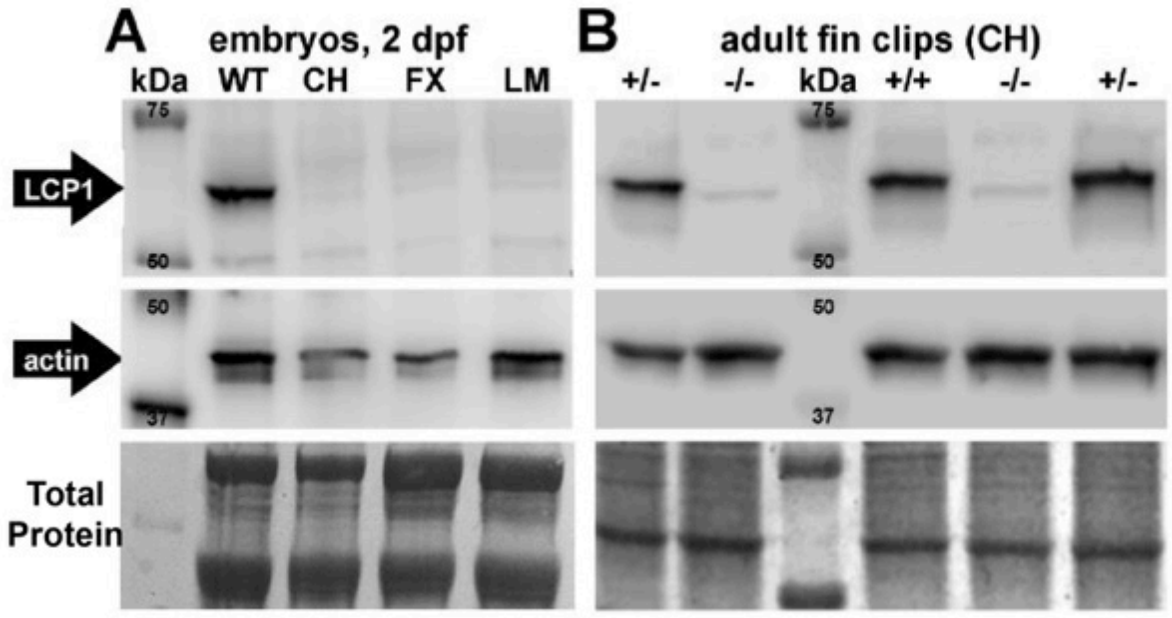

Figure 6: LCP1 protein is undetectable in null embryos and adults.

A) Chemiluminescent Western blot comparing 2 days post-fertilization (dpf) total protein lysates from a wild type line (WT) and three independent *lcp-1* null lines (CH, FX and LM). The top panel shows LCP1 signal (~65kDa) and the middle panel shows actin signal(~43kDa), a positive control. The bottom panel reflects the total protein in each lane (Coomassie Blue stain).

B) Chemiluminescent Western blot comparing total protein lysates from 5 adult animals of one line (CH): two heterozygotes (+/), one wild type (+/+), and two nulls (-/-) Image source: (Kell et al., 2018).

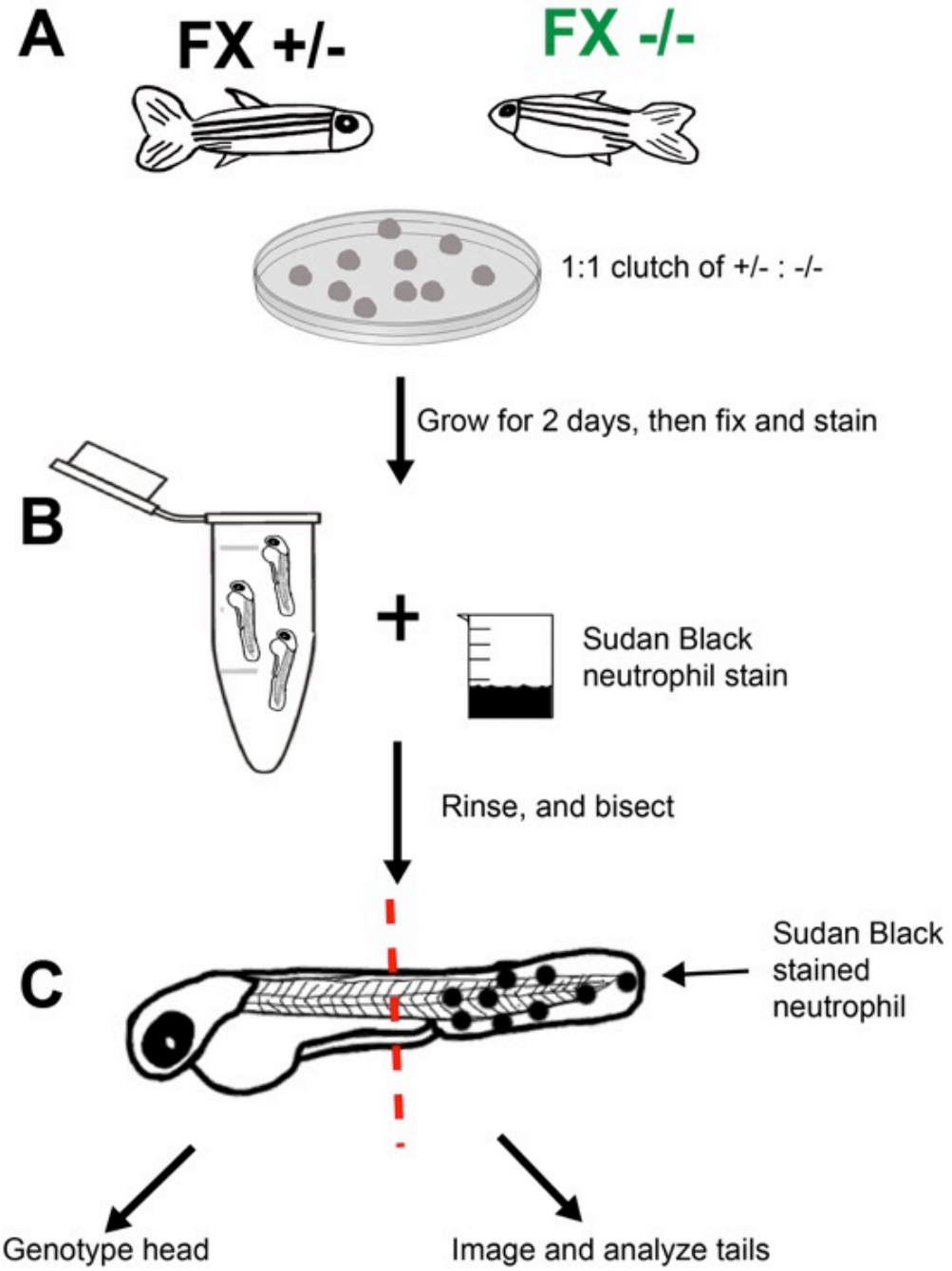

Figure 7: Procedure for staining larval neutrophils with Sudan Black.

A) FX heterozygotes (+/-) and homozygous mutants (-/-) are crossed to produce a clutch that is 1:1 (+/-) and (-/-) progeny. B) After development for 2 days, fish are euthanized, fixed and neutrophils stained with Sudan Black. C) Bisected fish are genotyped (head) and imaged (tail) to record the density of neutrophils in the caudal hematopoietic tissue (CHT).

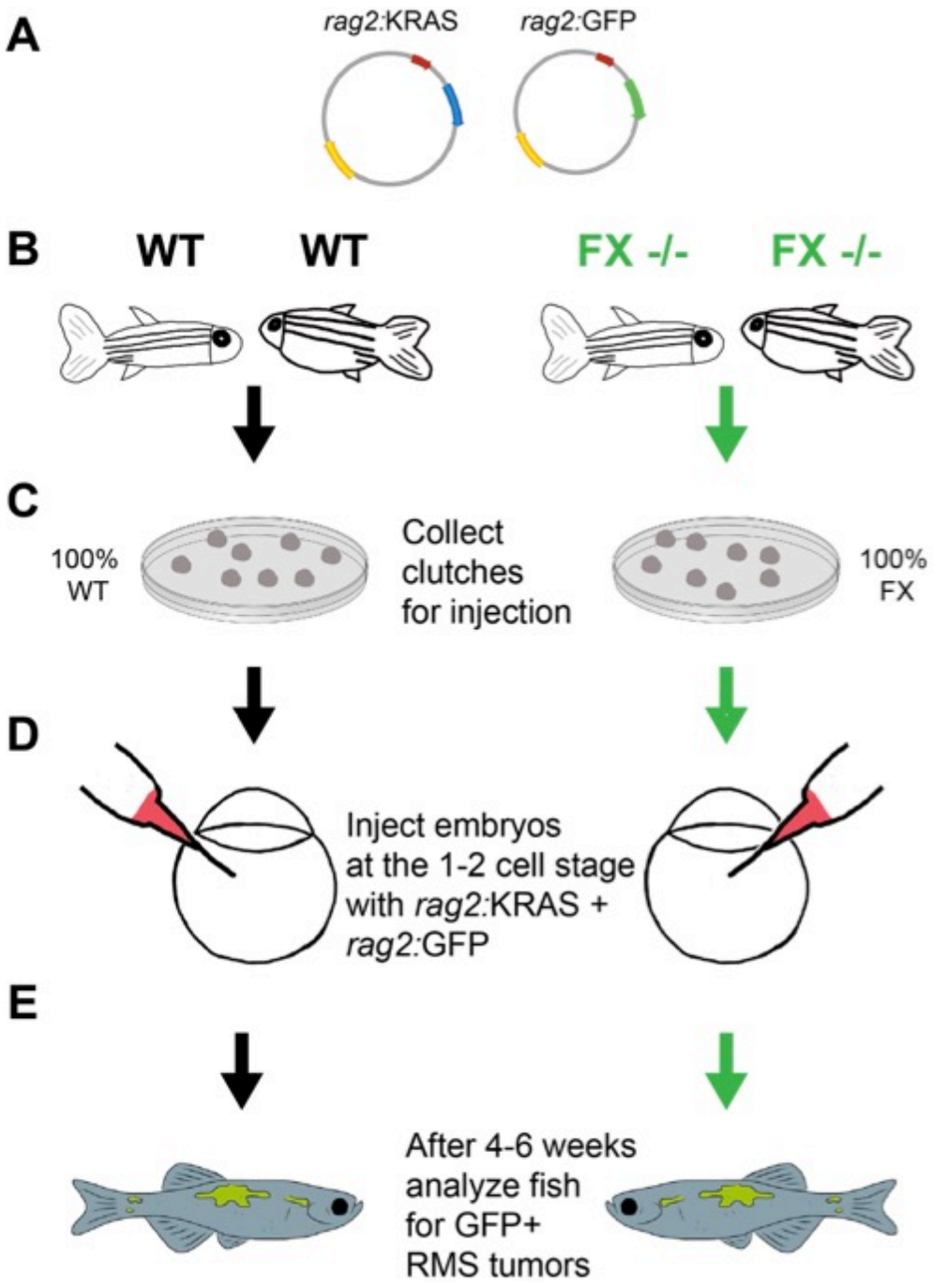

Figure 8: Procedure for inducing rhabdomyosarcomas (RMS) in zebrafish.

A) Brief plasmid maps of rag2:KRAS and rag2:GFP where red = the rag2 promoter site, yellow = ampicillin resistance, blue = the human KRAS oncogene and green = the GFP genes. B-C) Fish of the desired genotype (WT = wild type; FX = Foxtrot mutant) were crossed and offspring collected. D) At the 1-2 cell stage, zygotes were co-injected with one oncogenic plasmid (rag2:KRAS) and one reporter plasmid (rag2:GFP) . E) After 4-6 weeks of monitoring, some percentage of fish are predicted to develop RMS tumors that are GFP+.

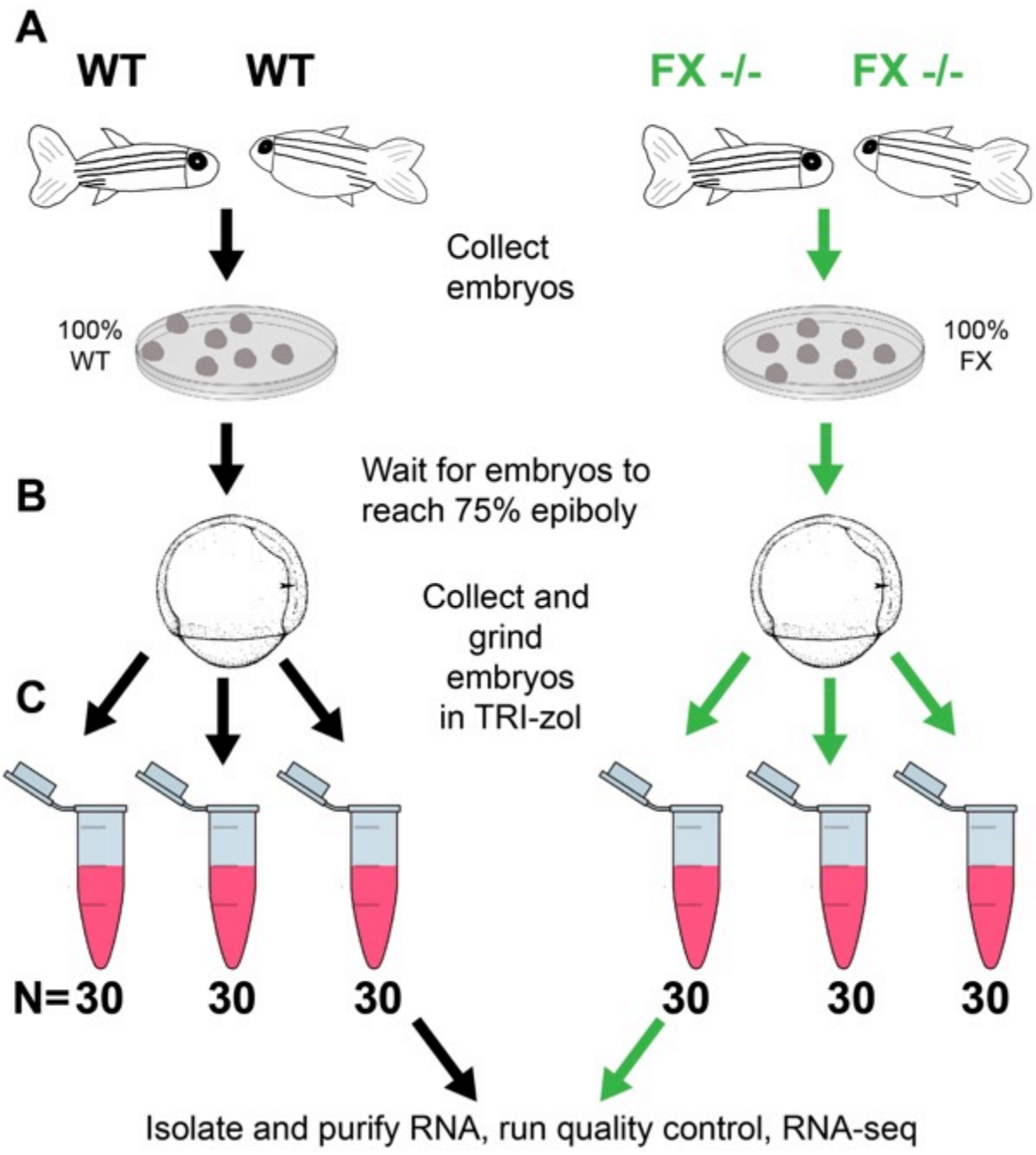

Figure 9: Procedure for analyzing the transciptome of wild type and L-plastin mutant embryos.

A) Fish of the desired genotype (WT = wild type; FX = Foxtrot mutant) were crossed and offspring collected. B) Once embryos reached 75% epiboly, three cohorts of 30 embryos per clutch were homogenized in TRI-zol reagent. C) Total RNA was evaluated by gel electrophoresis and spectrophometry, then submitted for RNA-seq analysis.

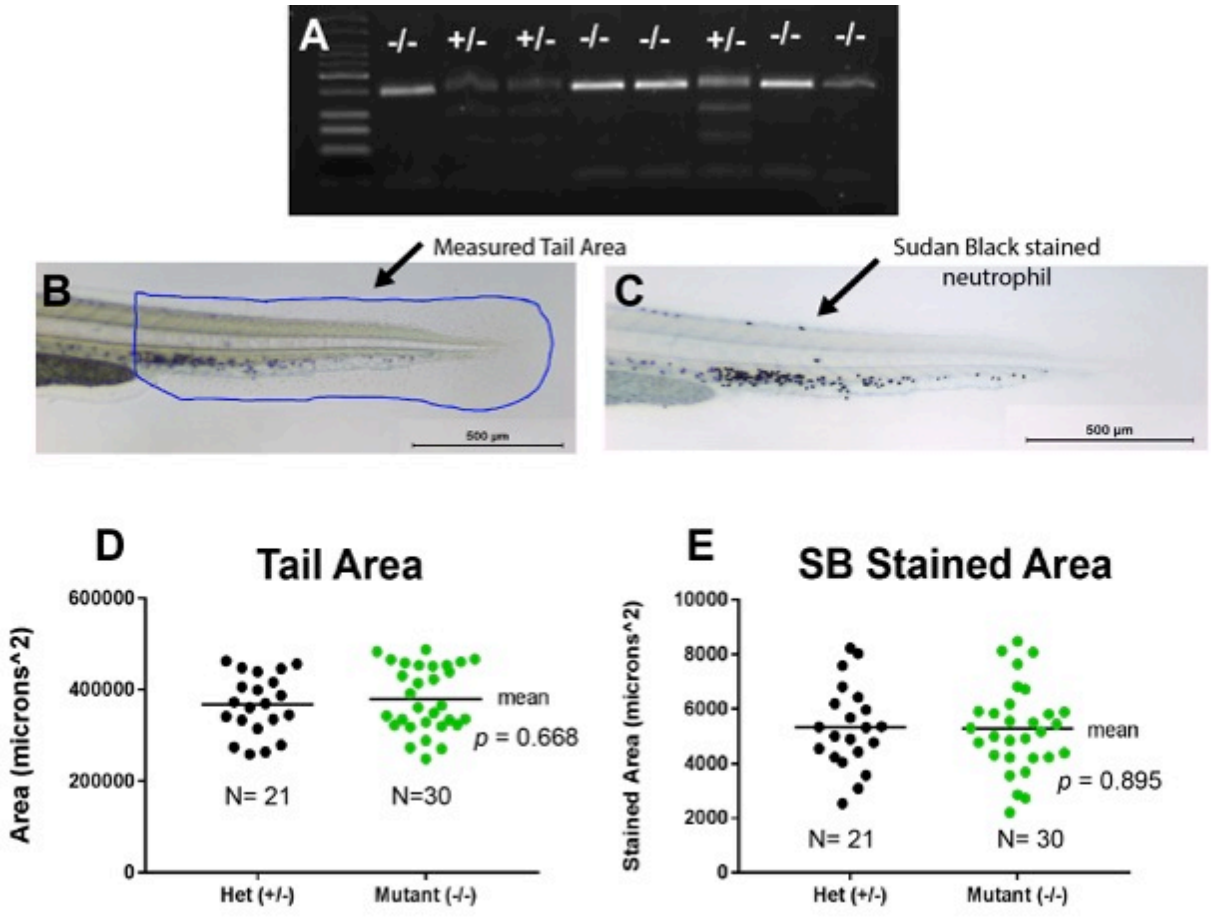

Figure 10. Neutrophil staining results.

A) Representative genotyping gel of 1 dpf embryos shows that the clutches are 50:50 for heterozygous (+/-) : homozygous mutant (-/-) genotypes

B) A tail area measurement in ImageJ. Blue line = area measured.

C). A Sudan Black-stained tail, showing the distinctly dark neutrophils.

D). Dot plot of the distribution of tail areas.

E) Dot plot of the distribution of Sudan Black stained areas.

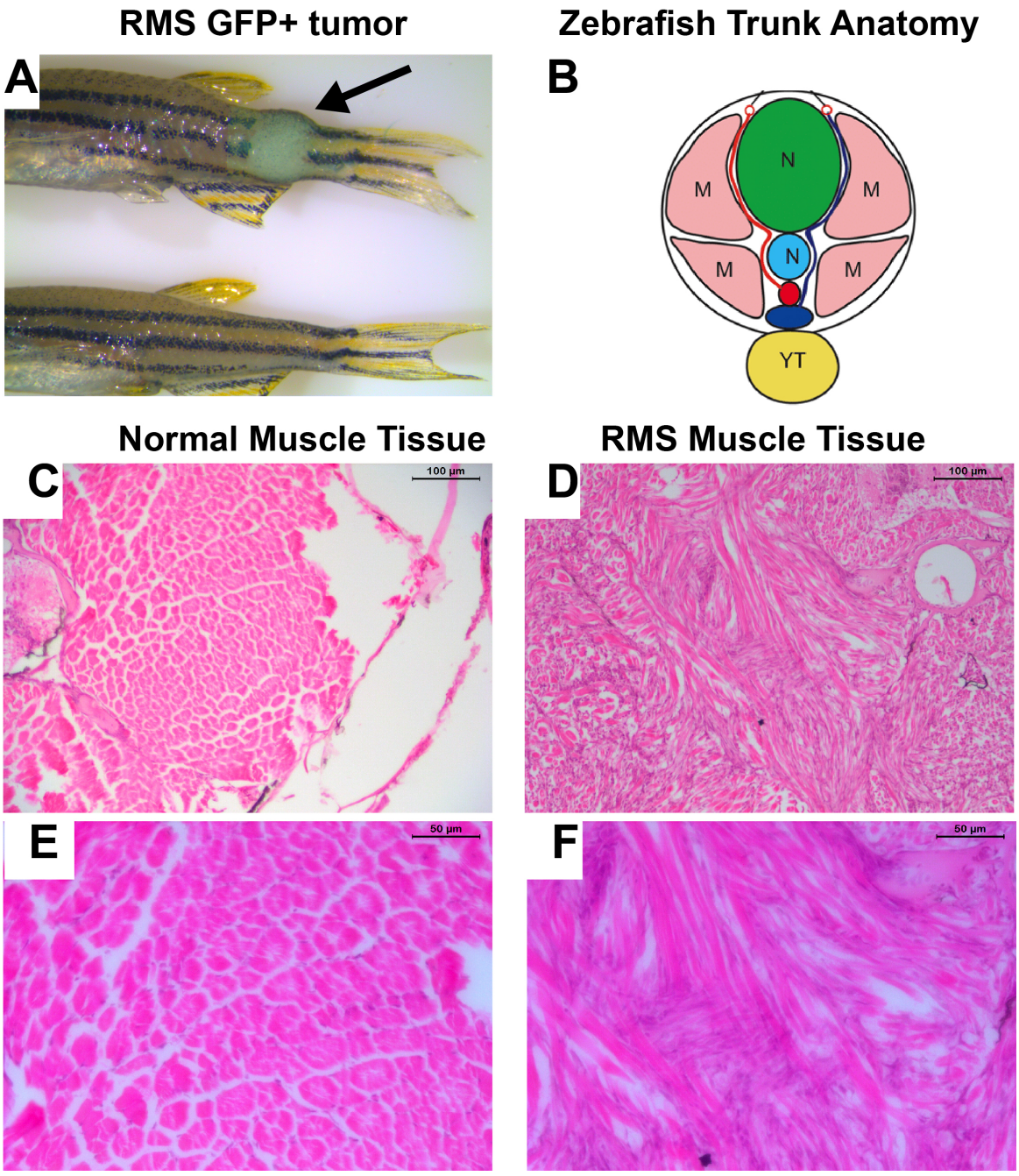

Figure 11. RMS tumor induction results.

A) Comparison of an RMS+ tumor fish and a normal sibling . The bright field image is overlaid with a fluorescent image of the same specimens. B) Cartoon of the anatomy of the zebrafish trunk. Longitudinal diagram of 48 hpf zebrafish embryo, with neural tube (NT, green), notochord (N, turquoise), posterior cardinal vein (PCV, dark blue), dorsal aorta (DA, red), intersegmental vessels (ISVs, dark blue and red according to origin), muscle blocks (M, pink) and yolk tube (YT, yellow). Image Source: doi: 10.1371/journal.pone.0007716. C-F) H&E-stained cross sections of zebrafish trunk tissue. Images were taken from roughly the same position as the black box in panel B. C&E. Normal muscle histology. D & F. RMS cell histology.

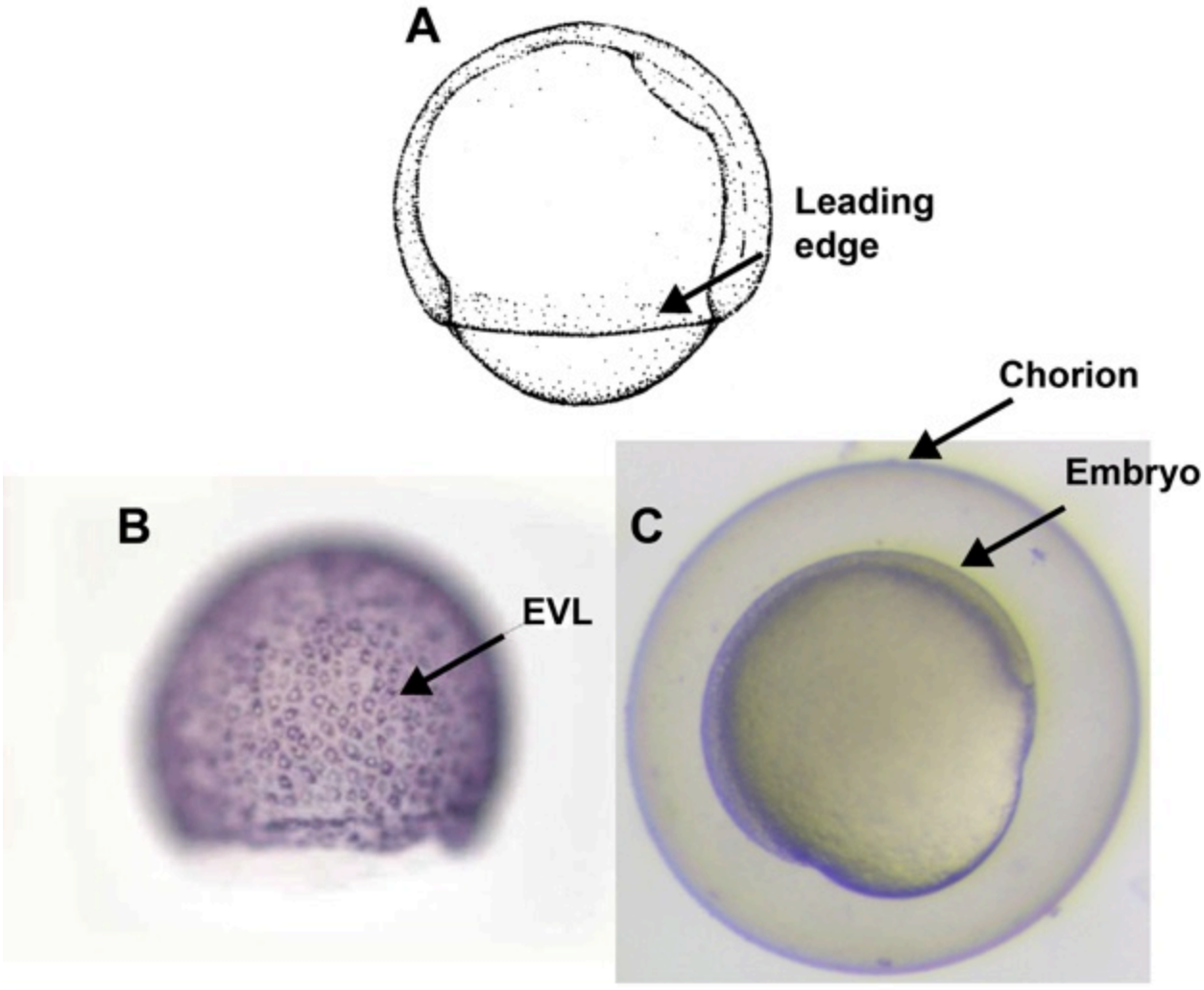

Figure 12. Staging of the zebrafish embryo at 75% epiboly.

A) Staging table diagram of 75% epiboly (Kimmel et al. 1995).

B) *in situ* hybridization for *lcp1* mRNA at the same stage. EVL = enveloping layer (Source: Thisse et al, ZFIN).

C) Light micrograph of a live embryo at 75% epiboly, prior to collection.

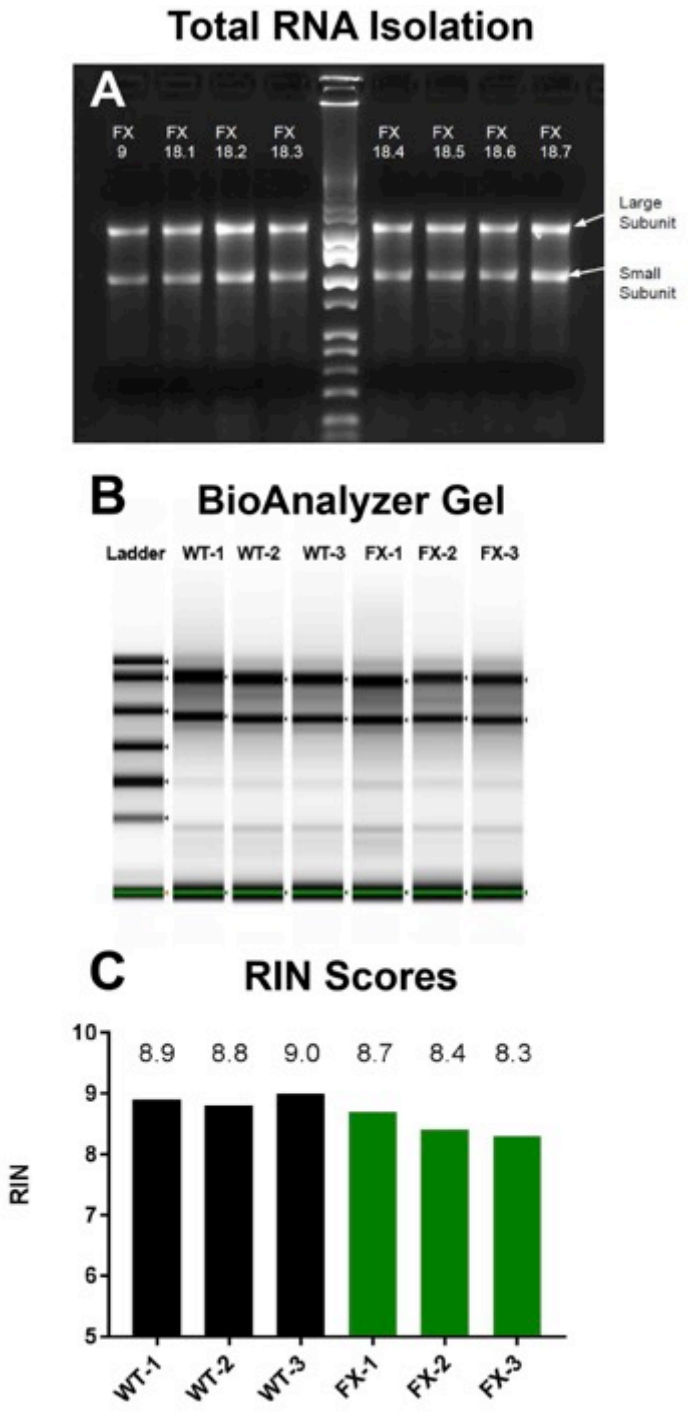

Figure 13. Quality control of zebrafish RNA samples.

A) An in-house RNA denaturing gel indicating that both large and small subunits of rRNA are present in roughly equal ratios, and that RNA has not degraded.
B) External quality control using the Agilent 2100 Bioanalyzer.
C) Bar graph of calculated RIN scores for each sample. For details, see Methods and Results.

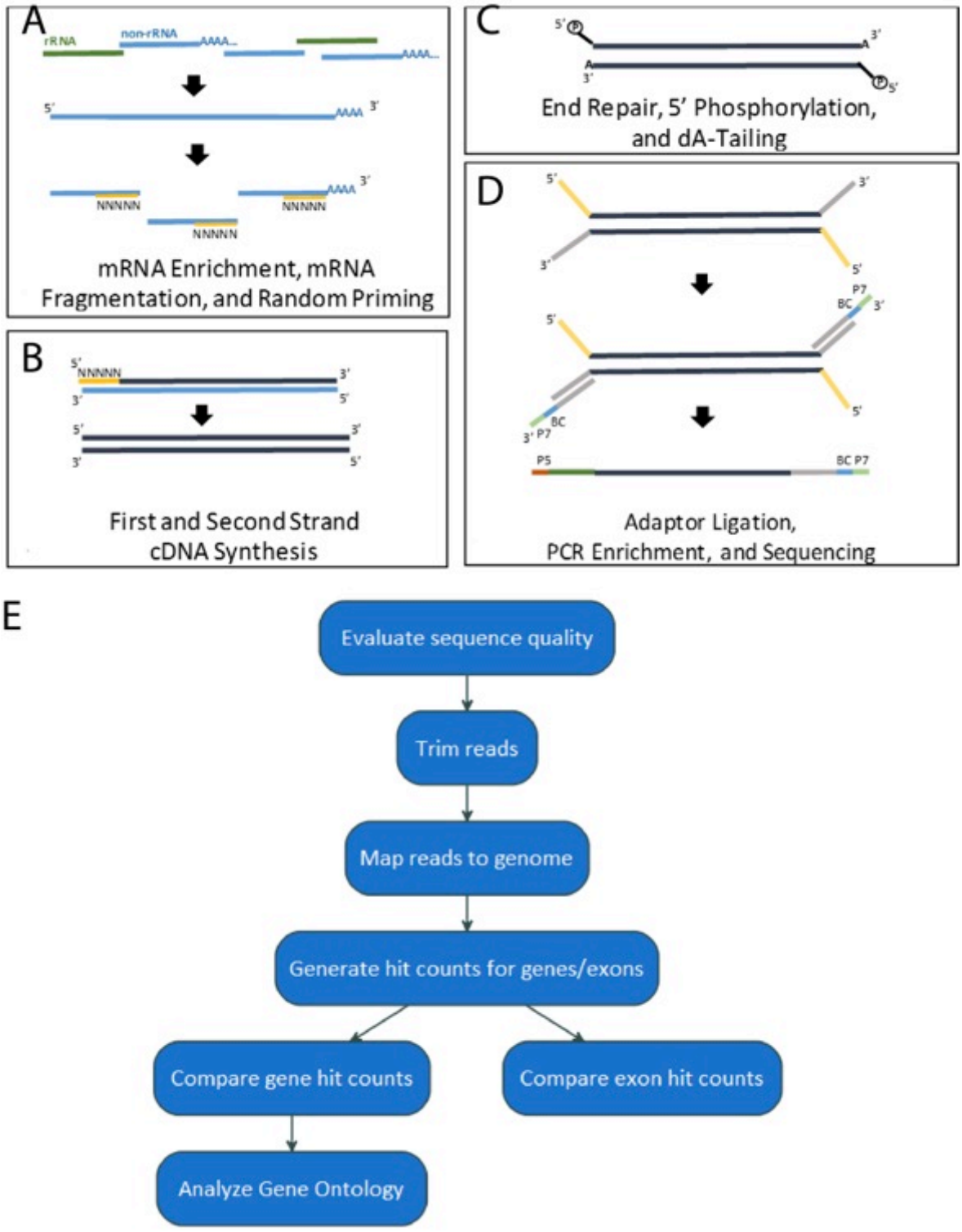

Figure 14. Workflow of RNA-seq analysis. (Source = Genewiz NGS)
A-D) Steps in cDNA library preparation prior to deep sequencing.
E) Steps in mapping and analysis after reads are sequenced.

Figure 15. Quality scores for the accuracy of base calls.
Each panel plots position of read on the X axis and quality score on the Y axis. The yellow points indicate the accuracy of the base assignments in each of the four samples processed for sequencing. All samples have high quality, within the green zone.
A) WT-1 B) WT-2. C) FX-1 D) FX-2.

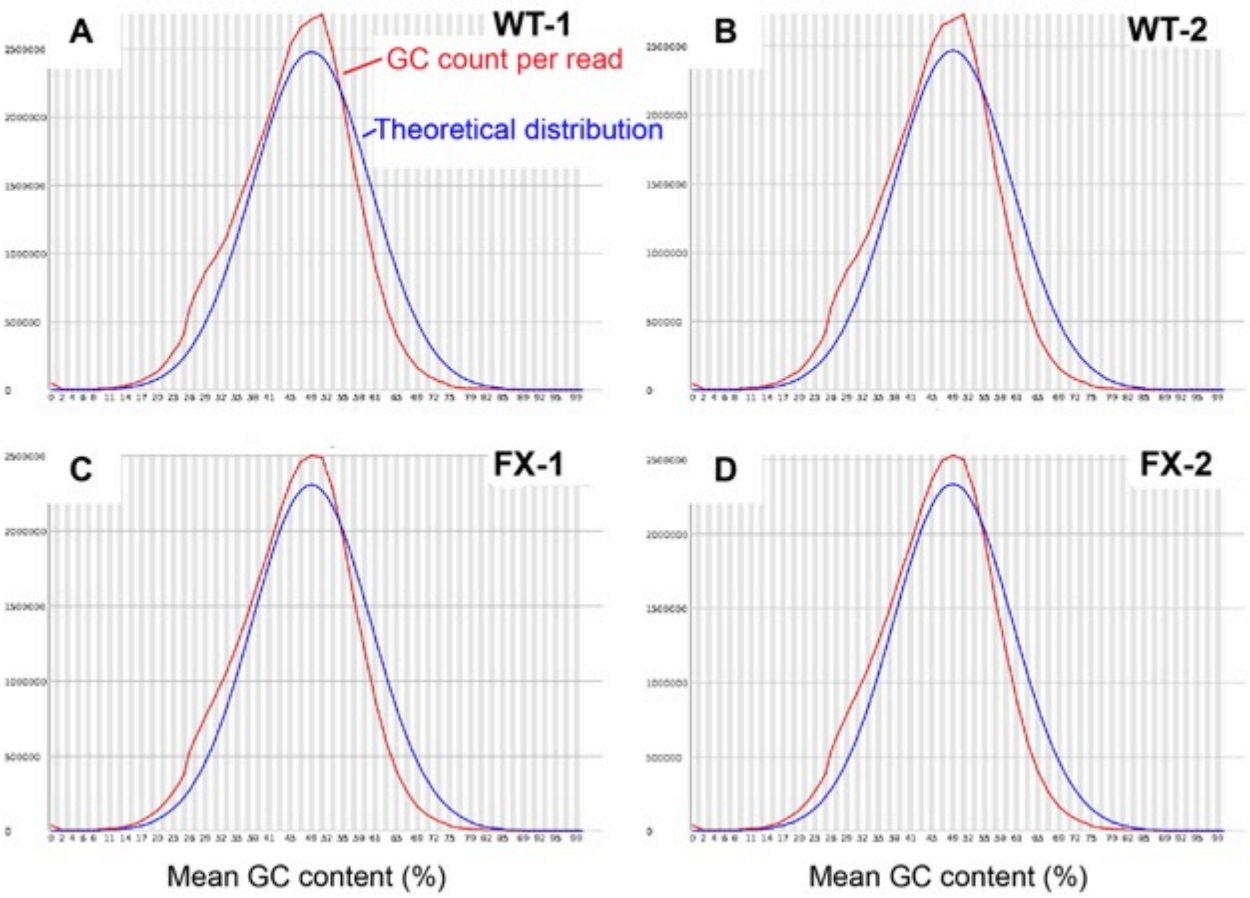

Figure 16. GC content of all samples.

Each panel plots percentage of CG content on the X-axis and number of reads for each percentage on the Y-axis. The blue line is the theoretical distribution for zebrafish and the red line is the actual distribution for each sample. All samples closely follow the theoretical standard bell curve of GC content.
A) WT-1 **B)** WT-2. **C)** FX-1 **D)** FX-2

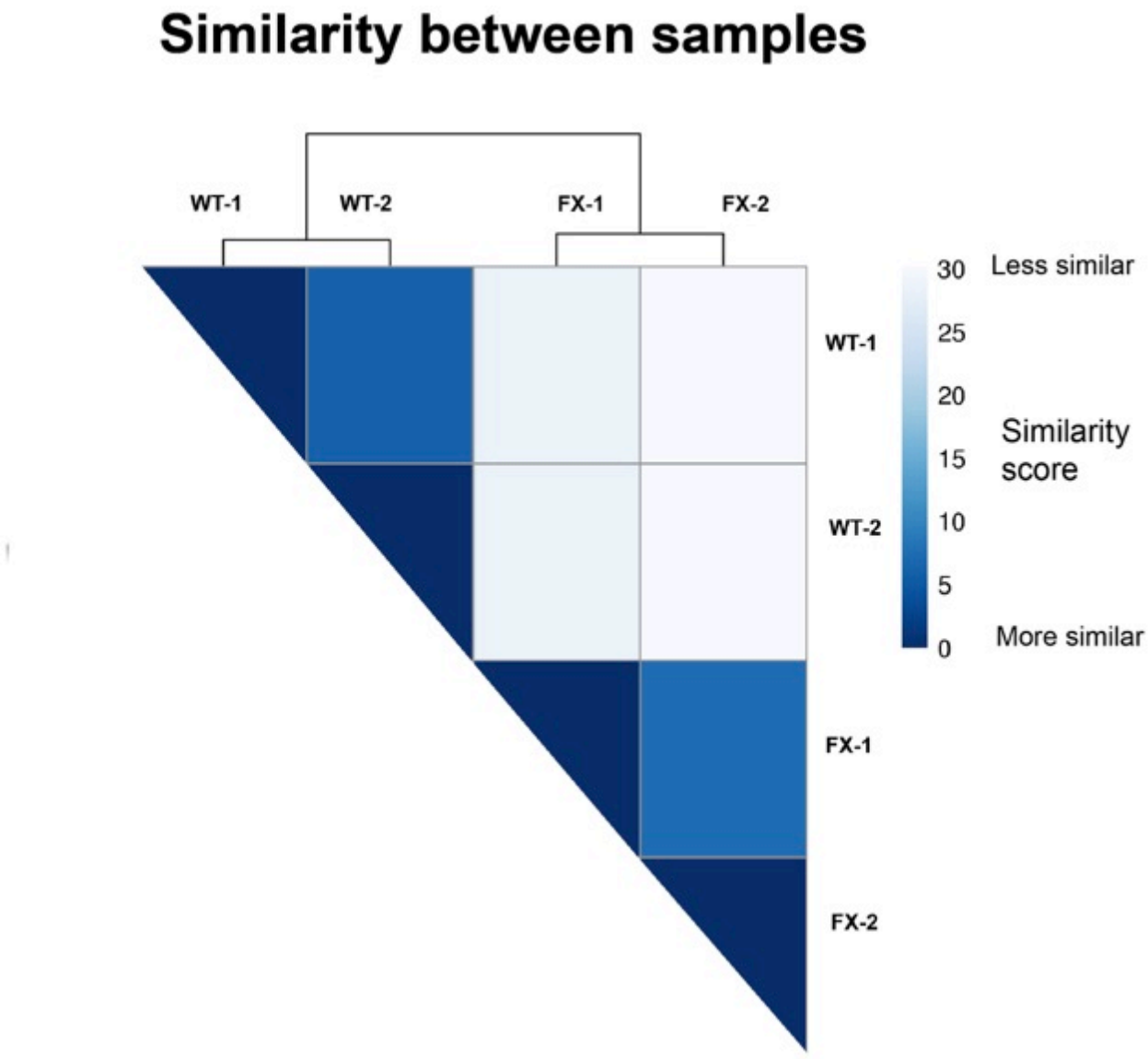

Figure 17. Transcriptomic similarity between samples.

Based on all mapped read counts after normalization, the darker the color, the more similar the samples. A score of zero (0) would indicate no differences and identical samples.
WT = wild type embryos. FX = 'Foxtrot' homozygous mutant embryos.

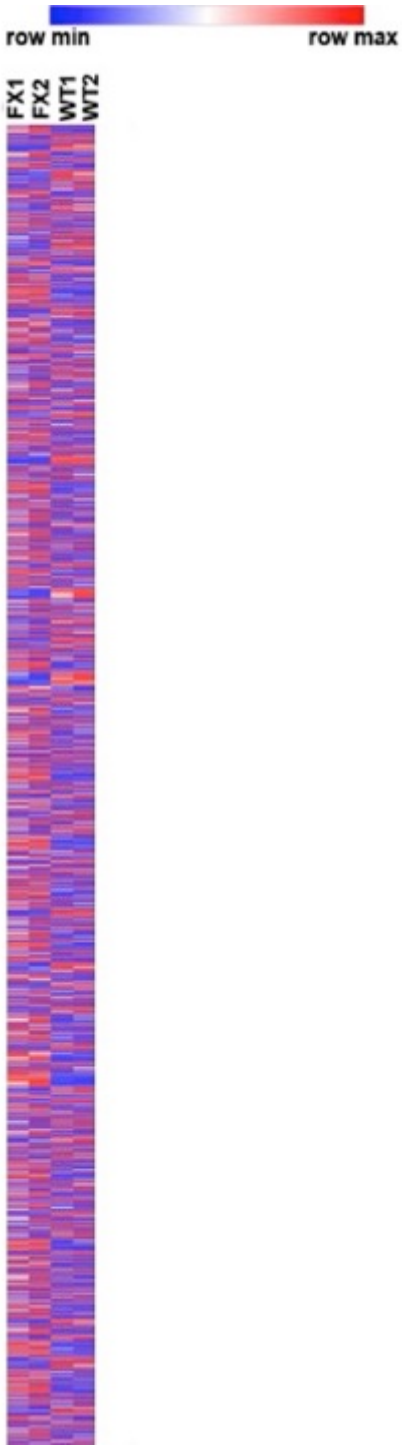

Figure 18. Heat map of differential gene expression for all four mRNA samples.

Each row represents a known genes in the zebrafish genome (N_{total} = 29292). Red indicates a high expression level, while blue indicates a lower expression level, relative to the average expression in each row based on normalized hit counts. Image created at: https://software.broadinstitute.org/morpheus/

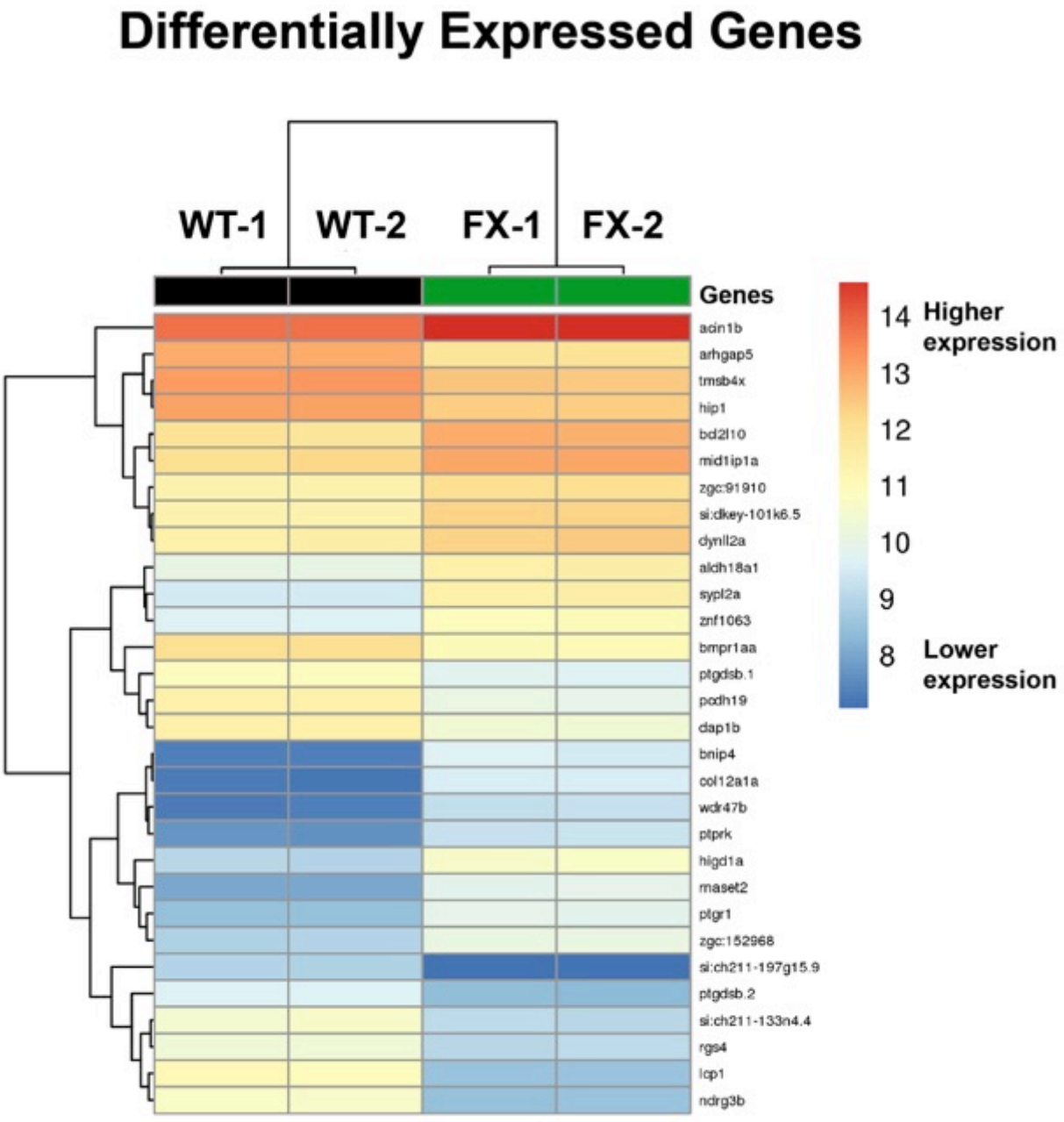

Figure 19. Heat map of the top 30 differentially expressed genes (DEGs) between samples.

A two-way cluster analysis of the top 30 differentially expressed genes (DEGs) based on adjusted p-values from DESeq2. Left to right, samples are clustered by overall similarity. Top to bottom, DEGs are clustered based on possible co-regulation. Darker blue indicates lower expression and bright red indicates high expression.

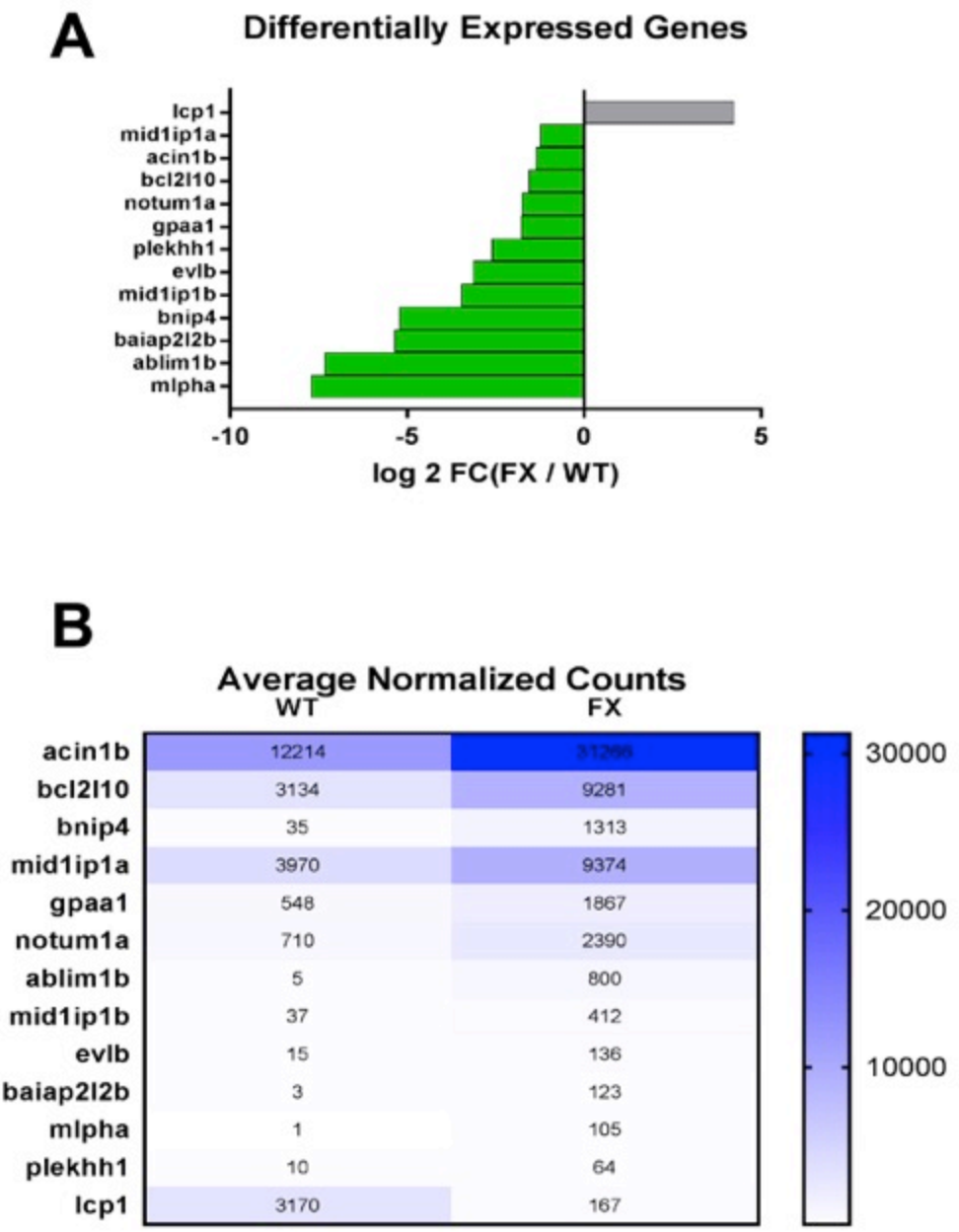

Figure 20. Detailed expression data for twelve genes of interest.

A) A visual representation of the log2 fold change (log2 FC) between wildtype and mutant samples for twelve genes of interest. Green bars = greater expression in mutant, Gray bar = greater expression in WT. B) Normalized count data averaged between the biological replicates for the same twelve genes. Dark blue indicates high-count transcripts (abundant); white indicates low-count transcripts (rare).

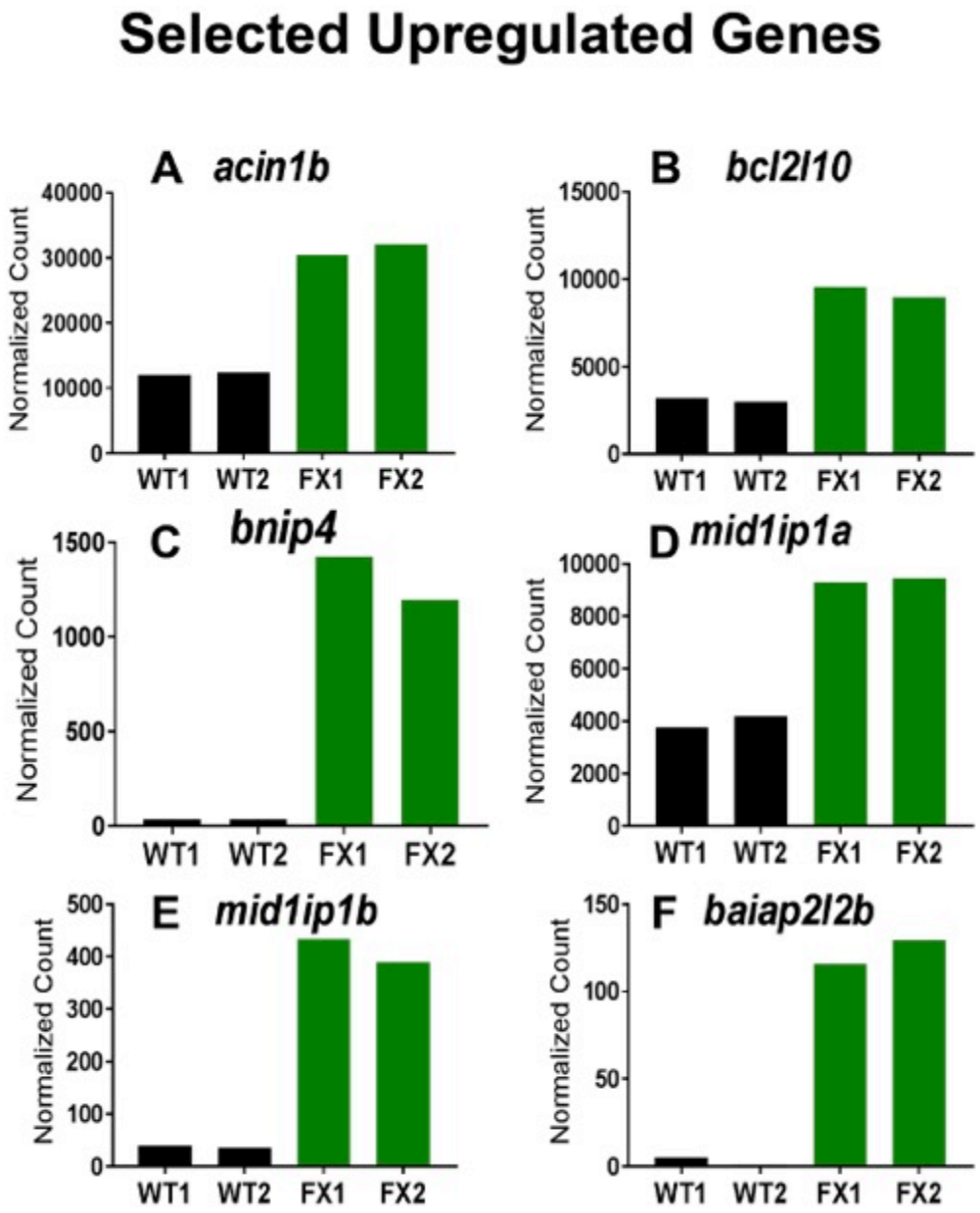

Figure 21. Bar graphs of normalized counts for six genes of interest. Bar graphs were created based on normalized hit counts for each sample and genes were chosen based on highly significant Wald statistics. From upper left to lower right, genes are ranked from abundant to rare. The complete gene names are given here:

acin1	apoptosis chromatin condenser inducer
bcl2l10	BCL2-like 10
bnip4	BCL2 interacting protein 4
mid1ip1a	MID1 interacting protein 1a
mid1ip1b	MID1 interacting protein 1b
baiap2l2b	BAI1-associated protein 2-like 2b

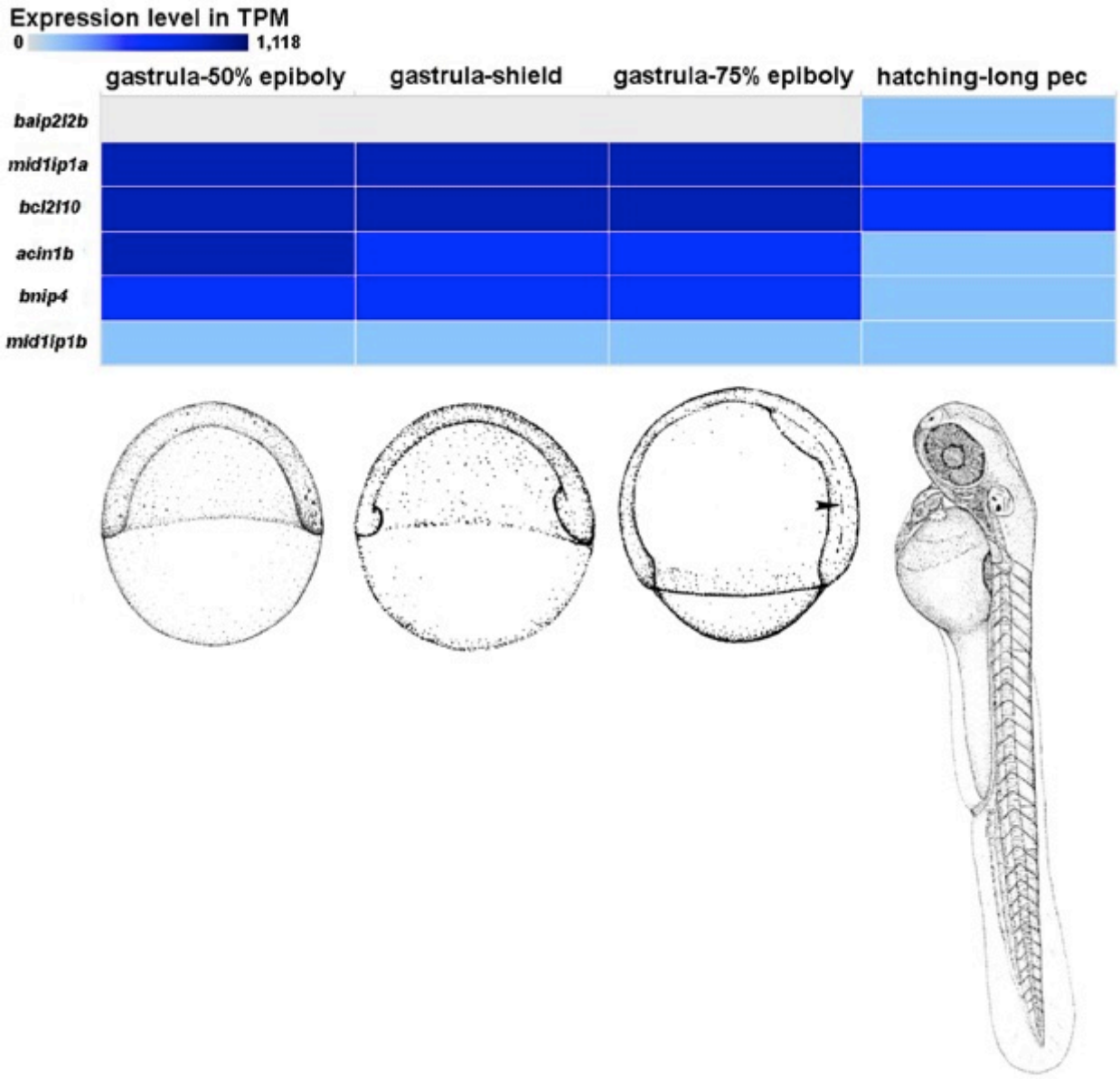

Figure 22. Known expression levels of specific candidate genes during zebrafish development.

WT expression levels of genes of interest over four developmental stages. The darker the color, the higher the expression level TPM = transcripts per million (Data Source: EMBL-EBI Expression Atlas).